The Dissertation Committee for Eva-Maria Strauch Certifies that this is the approved version of the following dissertation:

Characterization and Applications of the Twin-Arginine Transporter Pathway

Committee:

George Georgiou, Supervisor

Charles F. Earhart

David Hoffman

Edward Marcotte

Ian J. Molineux

Characterization and Applications of the Twin-Arginine Transporter Pathway

by

Eva-Maria Strauch, M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2007

Dedication

To my aunt, who supported me in every aspect of my life in the most selfless way. Without her I would not be where I am today. And, to my parents and sister, for the love they have shown me throughout my journey.

Characterization and Applications of the Twin-Arginine Transporter Pathway

Publication No._____

Eva-Maria Strauch, Ph. D.

The University of Texas at Austin, 2007

Supervisor: George Georgiou

The twin-arginine translocase allows the translocation of folded protein substrates across the cytoplasmic membrane of bacteria and archaea or the thylakoid membrane of plants. In *Escherichia coli*, its protein components TatA, TatB and TatC assemble dynamically upon interaction with protein substrates. Prior to export, the machinery performs a quality control check so that only correctly folded proteins are translocated. The first objective of this work was to derive and apply new methodologies based on the inherent qualities of the pathway.

We developed a new bacterial two-hybrid system that capitalizes on the folding quality control mechanism of the Tat pathway. One protein (prey) is fused to Tat-specific signal peptide. A second (bait) protein is produced as a fusion to a reporter that produces

a “signal” (growth or enzymatic activity) only when the bait-reporter fusion binds to the prey and the resulting complex is exported into the periplasm via the Tat pathway.

As a second biotechnological application of the Tat pathway, we developed a phage display system that allows the protein of interest to fold within the cytoplasm prior export and display onto phage particles. This is in contrast to the conventional phage display system, in which displayed protein folds in the periplasm. We took advantage of this new system to screen a library of 2×10^6 of fluorescent GFP variants containing a hexameric peptide insertion for ligand binding. Despite the diversity of the hexamer, we were not able to isolate single GFP variants that would bind with specificity to various ligands. This highlights the difficulty in engineering GFP variants that can bind to other proteins while retaining the ability to fluoresce.

The second aspect of this research was to examine mechanistic aspects of the Tat pathway. TatB and TatC are responsible for the recognition of Tat signal peptides. Here, we established the importance of TatC as the crucial component of the Tat pathway for the interaction with the hallmark twin-arginine motif within Tat signal peptide. Substitution of the RR dipeptide with a KK sequence completely abolishes export. In a genetic screen using a ssTorA(KK)-GFP-SsrA as a reporter. We identified several amino acid substitutions within TatC that allowed the alteration of the substrate specificity of the pathway as indicated by the impairment of indigenous Tat substrates. Finally, we analyzed the conformational dynamics of TatA using GFP fusions and by incorporation of the chemically reactive, non-canonical amino acid azidohomoalanine.

Table of Contents

List of Tables

List of Figures

Chapter 1

Protein translocation: molecular basis and biotechnological applications

1.1 INTRODUCTION

Every living organism contains a protecting lipid bilayer as a membrane which separates itself from the environment. One theory is that life could not have developed without the spontaneous self-organization of lipids and amphiphiles into supramolecular structures such as micelles and lipid-bilayers (Paleos et al. 2004). In living cells today, biological membranes have several roles. They are involved in compartmentalization and energy transduction, nutrition and ion transport, signal transduction and membrane-bound enzyme-catalyzed metabolic reactions.

1.2 ENERGY SUPPLY FOR PROTEIN TRANSLOCATION

Proteins are charged, bulky heteropolymers of which transport across or insertion into the low dielectric barrier of a membrane is thermodynamically highly unfavorable. Hence, there are several different transport pathways coupling the metabolic energy supplies to these processes to overcome this physical barrier. Most translocases utilize energy generated by the hydrolysis of nucleoside triphosphates, e.g. ATP or GTP. However, there are several different transport processes across biological membranes which solely rely on existing ion gradients.

In general, protein translocation can be summarized as a basic set of rules (Schatz and Dobberstein 1996). A precursor protein containing a targeting sequence which is typically N-terminal, is maintained in an unfolded state prior to export. This unfolded, export-competent state of the precursor protein is either achieved by cytoplasmic chaperones (post-translational) or by pausing translation followed by an immediate association of the protein-synthesizing ribosome with a receptor in the membrane (co-translational export). The actual protein translocation is typically mediated by a set of transmembrane proteins which form a hydrophilic channel.

1.2.1 The role of NTP in protein translocation

Nucleoside triphosphate (NTP) hydrolyzing protein machineries typically contribute the majority of the driving force for protein transport. NTP driven mechanisms are differentiated based on which side of the membrane they are active: *cis* (pushing force) or *trans* (pulling force). An example for a *cis* active motor is the bacterial SecA protein which utilizes ATP for the threading around 25 amino acid in an unfolded state through the membrane (van der Wolk et al. 1997). An example for a *trans*-acting protein is the chaperone Hsp70 in eukaryotic cells which facilitates transport into the endoplasmatic reticulum (ER) or the mitochondrial matrix. In general, the mechanism of action can be described as a processive moving-along the unfolded polypeptide chain. Energy derived from NTP hydrolysis results in pulling or pushing of the polypeptide chain thus facilitating transport.

1.2.2 Proton motive force

In addition to the energy derived from ATP or GTP hydrolysis cycles, most protein transport systems also use energy provided by the electrochemical gradient across the membrane. We are only starting to understand how the different transport systems adapt the electrochemical gradient into a driving force at a molecular level. For example, the energy stored in the mitochondrial inner membrane potential is absolutely necessary for the transport of matrix-targeted proteins even though the import machinery also involves Hsp70 and the use of ATP hydrolysis. This phenomenon applies similarly to the Sec machinery. The presence of the protein motive force (PMF) drives the translocation through the Sec pore once the polypeptide has been released from SecA. Further, the PMF prevents the backward slipping of the substrate (Schiebel et al. 1991; Driessen 1992). However, at least for the Sec pathway, the PMF cannot initiate translocation (Driessen 1992).

Energy stored in the proton motive force has two components: the electrical field gradient ($\Delta\psi$) and the pH gradient (ΔpH). The proton gradient is generated by the free energy supplied by a downhill electron transfer chain. The electron transport chain catalyzes the oxidation of certain donor substrates e.g. NADH by certain acceptor substrates e.g. oxygen and couples this reaction to proton translocation across the membrane. Typically, at least two topological separate enzymes are required, a dehydrogenase and a reductase. The two are coupled functionally to a quinone that diffuses freely through the membrane. The dehydrogenase catalyzes the reduction of the quinone by the donor substrate whereas reoxidation of the quinol is catalyzed by the reductase which transfers electrons to an appropriate substrate. Protons translocated across the membrane establish the electrochemical gradient, which the cell uses mainly

for the generation of ATP via the ATP synthase, the establishment of ion gradients ($\Delta\psi$), motility via flagella and directly or indirectly for protein translocation.

1.2.3 Breaking the rules

About 10 years ago, a remarkable pathway was discovered that exports proteins in a fundamentally different manner (Berks 1996). Protein transport via the Twin-Arginine Translocon (Tat) or the thylakoid ΔpH import pathway of plants does not follow the general sets of rules for protein translocation described in the section above. The Tat pathway solely relies on the membrane potential to export folded proteins across the cytoplasmic membrane. No NTP hydrolysis or *cis*-acting chaperones have been identified (Robinson and Bolhuis 2001). The precise role of the membrane gradient is unknown. For *E. coli*, the ΔpH is not required; however, $\Delta\psi$ is essential to drive Tat-mediated tanslocation (Bageshwar and Musser 2007).

1.3 PROTEIN TRANSPORT IN BACTERIA

All bacteria have a cytoplasm which is surrounded by a lipid-bilayer. In addition, Gram-positive bacteria are protected by a thick peptidoglycan layer whereas Gram-negative bacteria have a much thinner and differently structured peptidoglycan layer as well as an extra outer membrane, formed by an asymmetric lipid-bilayer. The space between the cytoplasmic membrane and the outer membrane is called the periplasm.

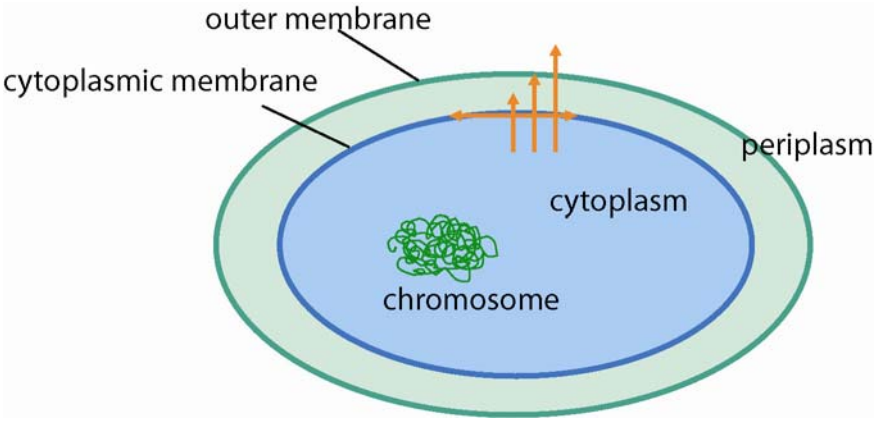


Figure 1.1. Overview of a Gram-negative bacterium and location choices for newly synthesized proteins. Indicated are the inner and outer membrane of the bacterium; the peptidoglycan layer which is located in the intermembrane space, has been omitted for simplicity.In orange arrows indicate the different routes a protein can follow.

In Gram-negative bacteria, a nascent protein exiting the ribosome can follow several different routes and end up at different final destinations. If the protein is not destined for the cytoplasm, it can be targeted into either the inner or the outer membrane, or across one or both membranes where it can localize in the periplasmic space or in the extracellular space respectively.

1.4 TRANSPORT SYSTEMS IN THE CYTOPLASMIC MEMBRANE

In gamma proteobacteria, the main route for protein transport across the cytoplasmic membrane is through the Sec translocon which is the bacterial homologue to the Sec61 in eukaryotic cells. From the perspective of the translocated protein, the main difference between the Sec and Tat translocation is the location of its folding maturation. For Sec substrates, folding and co-factor assembly occurs in the periplasm. Folding is

assisted by periplasmic chaperones such as DegP which functions as either protease or chaperone depending on the temperature, Skp, and four peptidyl-proline *cis-trans* isomerases, PpiA, PpiD, SurA and FkpA. The periplasm is an oxidizing environment mainly due to the action of DsbA which oxidizes free cysteines, whereas the isomerase DsbC rearranges disulfide bridges to their native conformation (Georgiou and Segatori 2005).

In contrast, Tat proteins fold within the cytoplasm. The cytoplasmic folding environment contains the general chaperones GroEL, DnaK/DnaJ/GrpE, the trigger factor, ClpB and the small heat-shock chaperones (IbpAB) (de Marco 2007) among others. Further for the incorporation of some cofactors, whole maturation pathways involving several enzymes are necessary. For example the incorporation of iron-sulfur cluster requires at least 8 proteins to ensure Fe-S incorporation (Tokumoto et al. 2002).

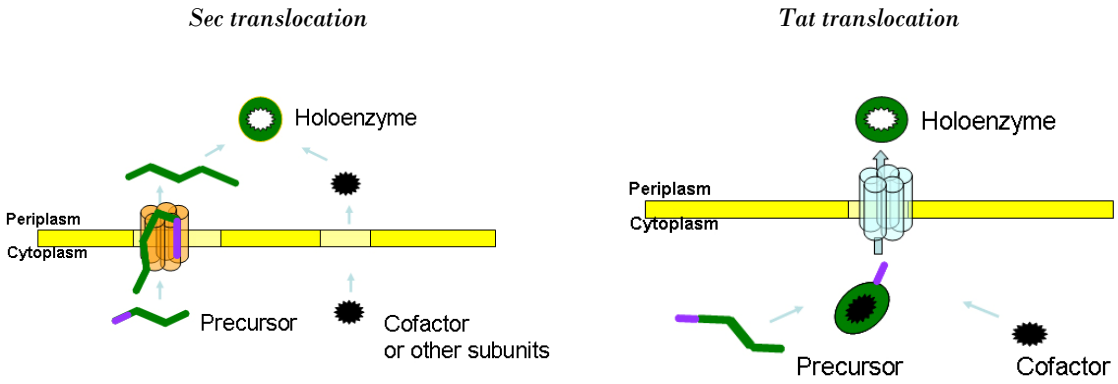


Figure 1.2. Protein export across the cytoplasmic membrane. The Sec pathway sequentially threads its substrates through the Sec pore and assembly occurs in the periplasmic environment whereas the Tat pathway only exports its substrates once they are fully assembled.

1.4.1 The Sec pathway

The Sec apparatus transports its substrates in an unfolded state through a narrow pore minimally established by SecYEG. Translocation can occur in two different manners which have been referred to as the co-translational and the post-translational export mode. Co-translational export involves the signal recognition particle (SRP) which is composed of the protein Ffh (fifty four homologue, based on its similarity to the eukaryotic SRP version in the endoplasmatic reticulum) and a 4.5S RNA unit (Luirink et al. 1992). The co-translational mode is often also referred to as the SRP pathway; it is ubiquitously found in all three kingdoms of life, however, *E. coli* operates the simplest version.

SRP binds to the signal peptide or a highly hydrophobic peptide stretch that exits the ribosome as a nascent polypeptide chain (Kim et al. 2001) (Fig.1.3.A). The loaded SRP complex binds to the membrane-bound receptor FtsY which mediates the interaction of Ffh with the Sec translocon (Fig.1.3.B); GTP hydrolysis is necessary to release the signal peptide or a hydrophobic patch in case of a membrane domain, which allows the handing-over of the substrate to the Sec translocon. The ribosome is recruited to the membrane where the rest of the polypeptide is directly synthesized (see Figure 1.3.C). Proteins following the co-translocational export route are typically inner membrane proteins. There is only one known exception in *E. coli* in which a non-membrane protein is exported via the SRP pathway: DsbA. DsbA probably folds to rapidly to be maintained in an unfolded state which is required for post-translational Sec transport(Schierle et al.

2003). It is conceivable that the co-translational route could be utilized for any protein which is otherwise prone to aggregation in the cytoplasm.

The post-translational mode involves the tetrameric chaperone SecB which binds to a nascent chain exiting the ribosome to prevent its immediate folding (Fig.1.3.E). In SecB mutants other general chaperones such as GroEL and/or DnaK can compensate for the loss of SecB (Kumamoto 1991). The association with SecB maintains the protein in a transport-compatible state since only the unfolded protein can be threaded through the membrane. SecB typically transfers its substrate directly to SecA. SecB binds dimeric SecA in an asymmetric fashion which presumably allows the dissociation of one SecA monomer, introducing a conformational change within SecA (Randall et al. 2005). A partial dissociation may allow the transfer step of the precursor protein to SecA. SecB eventually releases the precursor since it is not necessary for the translocation event (Fekkes et al. 1997). The translocation through the Sec pore is a step-wise event in which ATP hydrolysis by SecA allows the threading of around 20-30 amino acids of the polypeptide through the membrane (van der Wolk et al. 1997), (Figure 1.3.F).

1.4.2 The YidC pathway

YidC is a 60 kD transmembrane protein of *E coli*, which is essential for the survival of the cell. It is localized in the inner membrane and is the only export machinery that has been conserved in the inner membrane of mitochondria throughout evolution. It functions as a membrane protein integrase (Samuelson et al. 2000). In bacteria, it can either operate independently or associate with the SecDFYajC proteins of the Sec pathway assisting the lateral integration of transmembrane domains into the lipid bilayer (Economou and Wickner 1994). The YidC-assisted lateral transfer of the

hydrophobic domains of transmembrane α-helices into the bilayer prevents the Sec translocon from jamming (Luirink et al. 2005) (Figure 1.3.G). Only a few protein substrates that follow the Sec-independent mode of the YidC pathway have been identified (Figure 1.3.H). These are simple inner membrane proteins (IMPs), including the small phage coat proteins M13 and Pf3, and the endogenous IMPs F_oc of the ATP synthase (Facey and Kuhn 2004).

Recently, a third mode for a YidC dependent transport was discovered. The mechano-sensitive channel protein MscL is integrated into the inner membrane via YidC in a Sec-independent manner. In the case of MscL, the SRP promotes its membrane targeting via the FtsY receptor. Once associated with the membrane, MscL is integrated into the membrane by YidC solely (Facey et al. 2007) (Figure 1.3.D).

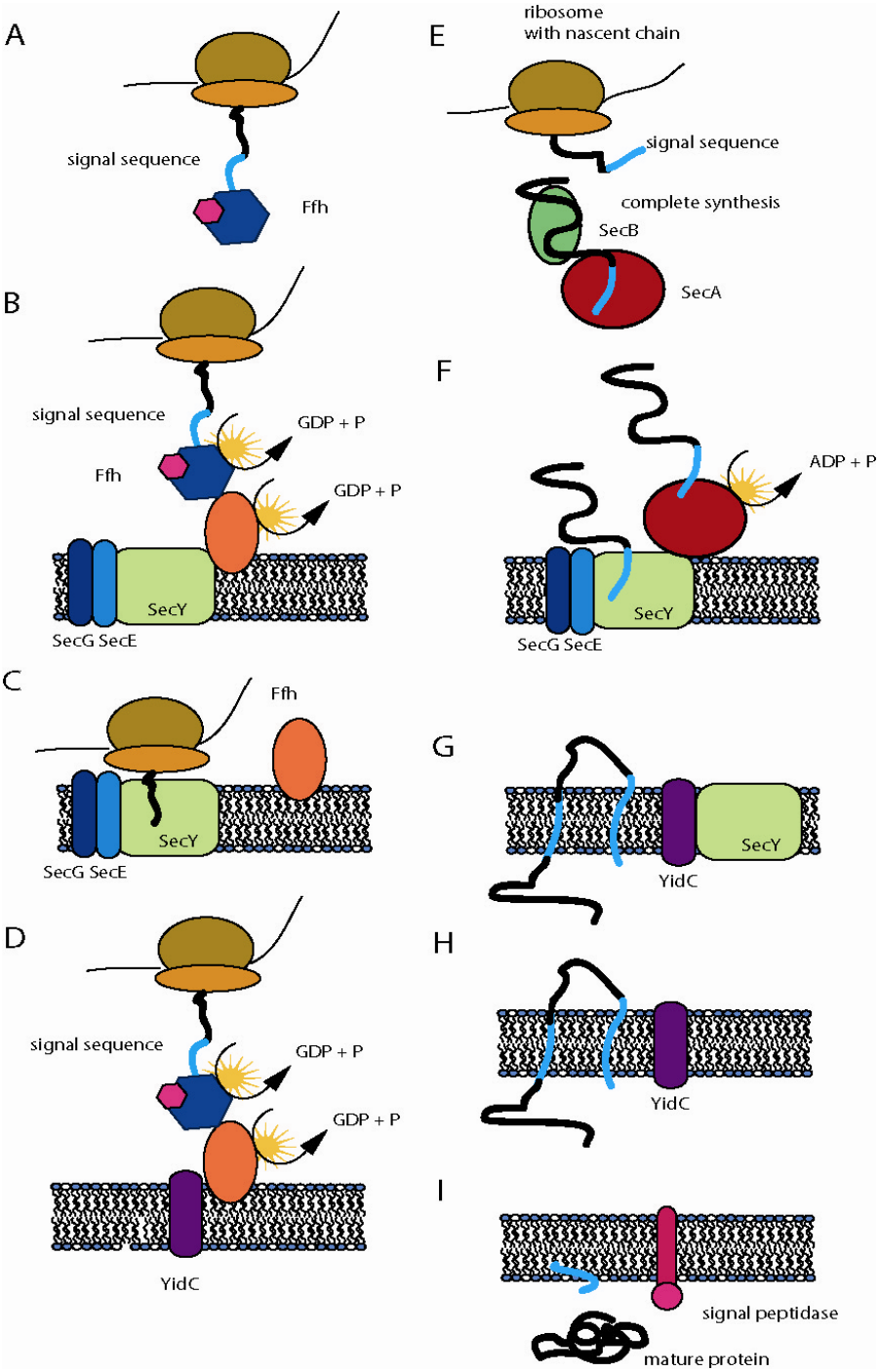


Figure 1.3. Overview of Sec- and YidC-mediated protein translocation.

1.4.3 The Tat pathway

Unlike the Sec pathway, the Twin-Arginine Transporter pathway exports proteins in a folded state through the cytoplasmic membrane. The minimal composition of the Tat translocon consists of the protein components TatA, TatB and TatC (Figure 1.4). TatB is dispensable in some Gram-positive bacteria or can replaced by mutated TatA variants (Blaudeck et al. 2005). In early studies, TatE, which is 50% identical to TatA, was thought to be an essential component of the translocon. However, TatE is not localized within the *tat* operon, it is 50-100 times less expressed than TatA (Jack et al. 2001) and its deletion has no impact on the export of model protein substrates (Sargent et al. 1998). Hence, all functional Tat components are encoded within the *tat* operon (*tatA,B,C,D*) which also contains TatD, a protein containing DNAse activity which is not involved in in Tat mediated transport (Wexler et al. 2000).

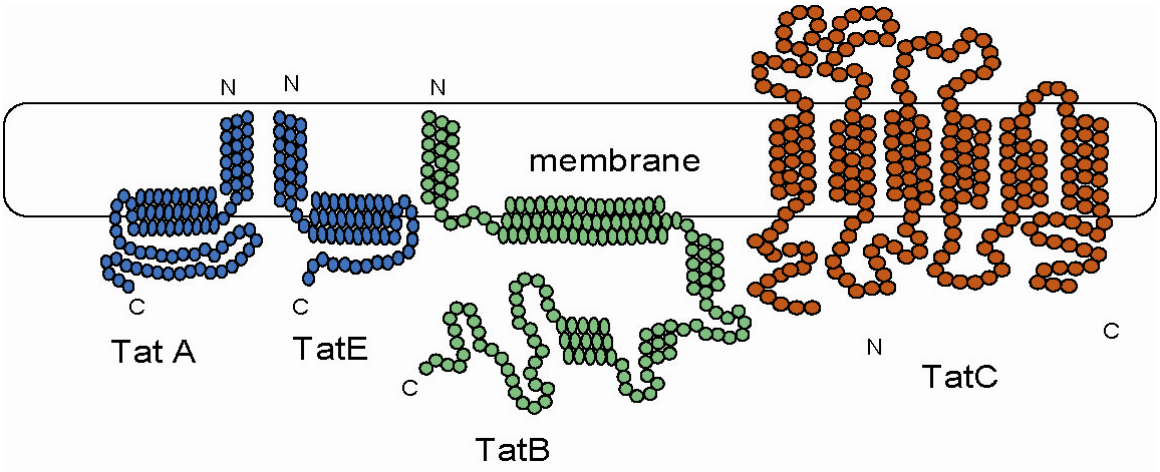


Figure 1.4. Protein components of the twin-arginine translocon and possible topologies.

1.5 THE TAT PATHWAY

The Tat pathway was discovered in bacteria around 10 years ago (Berks 1996), and in the thylakoids membrane of plants 15 years ago (Cline et al. 1992). Little is known about the detailed mechanism of protein translocation via the Tat pore. The most remarkable feature of the Tat pathway is that it exports completely folded and assembled substrates that are monitored for correctness of folding by an inherent proof-reading mechanism (DeLisa et al. 2003). The pathway solely operates on a proton motive force (Alder and Theg 2003; Bageshwar and Musser 2007) and so far no ATP requirement was demonstrated. The Tat pathway is highly conserved in archaea, in most bacteria and in the chloroplasts of plants. However it is not presereved in the mitochondria of most eukaryotes. Only a few protozoa encode Tat components in their mitochondrial genome. For example, the protist of Jakobis family *Reclinomonas americana* has the largest and eubacteria-like mitochondrial genome known to date (Lang et al. 1997). This organism carries next to a TatC homologue a TatA homologue in its mitochondrial genome (Gray et al. 2004).

1.5.1 Tat components

1.5.1.1 TatA

TatA is the most abundant component of the Tat export apparatus and it is believed to mediate the actual translocation step. Originally, TatA was postulated to contain one transmembrane domain with its N-terminal end facing the periplasmic side of the cytoplasmic membrane, and a second amphipathic helix located within the cytoplasm,

but associated with the lipid bilayer (Porcelli et al. 2002). TatA forms both *in vivo* as well as when solubilized in detergent oligomeric complexes. When surveying purified TatA, asymmetric pore-like formations with lid-like structures can be detected by single particle electron microscopy (Gohlke et al. 2005), see Figure 1.5.

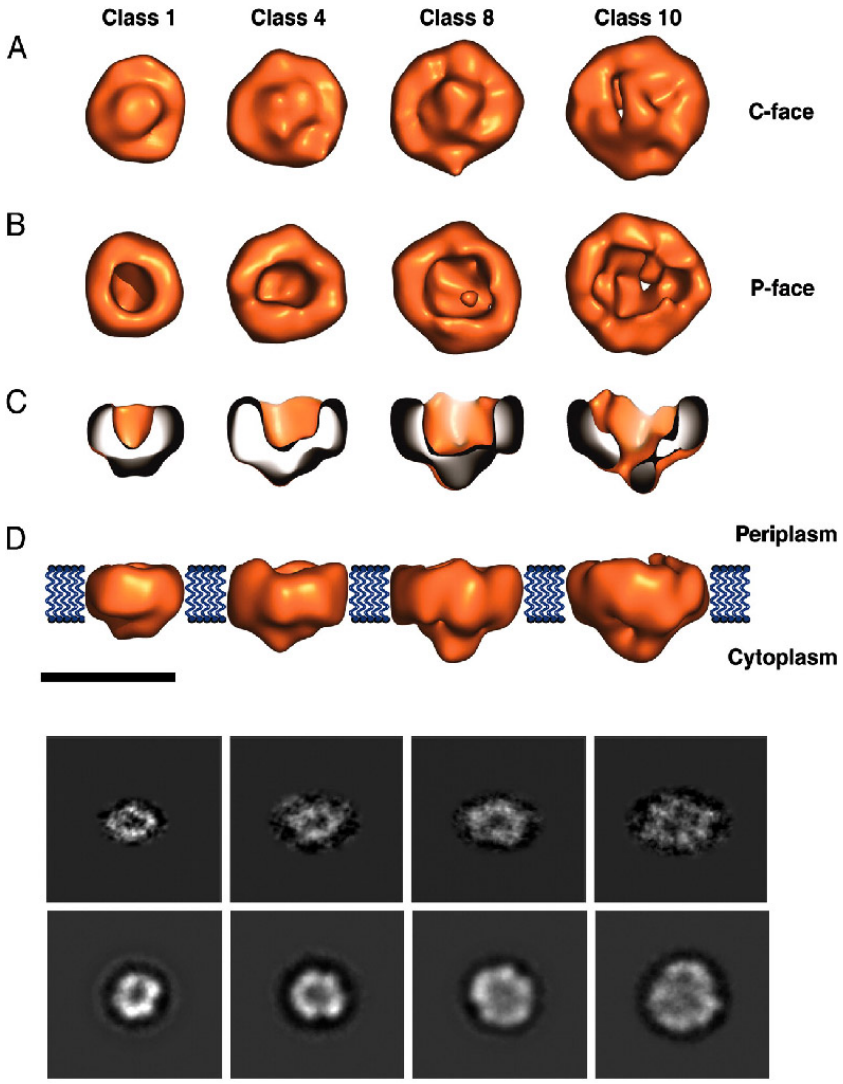


Figure 1.5. Single particle electron microscopy of purified TatA in detergent. The size and the pore-like formation indicated that these oligomers might accommodate transport of Tat substrates. Class 10 oligomers contain about 30 TatA oligomers and are 15 nm in width with a possible 8 nm diameter for a pore. Adopted from (Gohlke et al. 2005) Copyright © Proc. Natl. Acad. Sci. USA

Two independent topology studies claim two different orientations of TatA, but both were in agreement about a dual topology of the amphipathic helix of TatA.

When the C-terminal end of TatA was fused to the periplasmic reporter alkaline phosphatase (AP), AP activity was demonstrated suggesting a periplasmic location for the C-terminus. On the other hand, fusion of TatA to the cytoplasmic reporter β-glucuronidase (UidA) indicated a cytoplasmic location for the C-terminus. This evidence points towards a possible dual-topology of TatA with its N-terminal end facing the periplamic site but with an ambiguous location of the amphipathic helix which can be either periplamic or cytoplasmic (Gouffi et al. 2004)(see Figure 1.6.A).

Chan and co-workers (Chan et al. 2007) proposed an alternative model for the topology of TatA that contradicted the above studies (Figure 1.6.B). Utilizing a membrane permeable and a non-permeable thiol-reactive agent, they demonstrated that a cysteine residue introduced within the N-terminus at position 2 reacted only with the membrane permeable probe indicating a cytoplasmic localization of the N-terminus. A cysteine residue introduced at the C-terminal end of the transmembrane domain on the other hand was labeled by the impermeable probe, indicating accessibility from the periplasmic site (Chan et al. 2007). Further, they claimed that the amphiphatic helix can be found either within the membrane or within the cytoplasm when using the same detection method.

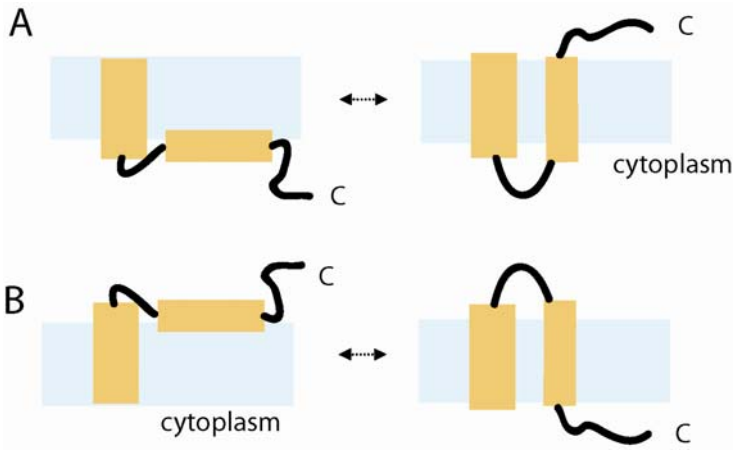


Figure 1.6. Models for the dual topology of TatA. The dual topology of the C-terminus has been confirmed by several different studies. The location of the N-terminus remains controversial.

1.5.1.2 TatB

The protein TatB is the least-studied component of the Tat machinery. For *E. coli*, and most other Eubacteria and Archaea, TatB is an essential component of the Tat translocon (Bogsch et al. 1998). However, a few Gram-positive organisms do not encode any TatB homologues. Hence, it has been proposed that TatA might have a bi-functional role when TatB is missing (Blaudeck et al. 2005; Barrett et al. 2007). A further possibility is that only certain substrates require TatB for translocation. Even in *E. coli*, a few substrates such as the reporter protein MBP, can be exported via Tat without TatB, indicating a possible redundancy in the translocation event. It was also hypothesized that TatB might have a specific role in the inherent proof-reading mechanism of the Tat pathway which has not been confirmed yet (Fisher and DeLisa 2004).

TatB is 25% identical with TatA at the amino acid level and seems to have a similar topology as TatA, containing one transmembrane domain and one amphipathic helix,. The N-terminus of TatB forms the transmembrane domain, whereas the C-

terminus is aligned with the membrane on the cytoplasmic site. When the transmembrane domain of TatB is removed, the ampipathic helix is still associated with the membrane whereas the amphiphatic helix of TatA can only be found in the soluble fraction (De Leeuw et al. 2001). Truncation analyses demonstrated that only the first 101 N-terminal amino acids of TatB are necessary for function (Lee et al. 2002). TatB associates with TatC in an equimolar ratio in complexes of around 400 – 550 kD (McDevitt et al. 2006). As mentioned below the TatB/TatC complex is responsible for signal peptide recognition. TatB and TatC within the receptor complex exhibit distinct types of interaction with the signal peptide: TatC interacts with the twin-arginine motif (Alami et al. 2003; Kreutzenbeck et al. 2007b; Strauch and Georgiou 2007b) whereas TatB seems to interact with the hydrophobic part of the signal peptide (Alami et al. 2003).

1.5.1.3 TatC

TatC is the most conserved component of the Tat pathway. It has 6 transmembrane domains in *E. coli* with both the C- and N-terminus facing the cytoplasmic side of the inner membrane (Buchanan et al. 2002; Ki et al. 2004). It establishes, together with TatB, the signal receptor complex which specifically recognizes the twin-arginine motif (Kreutzenbeck et al. 2007b; Strauch and Georgiou 2007b). The function of TatC is discussed in detail in chapter 3.

1.5.2 Tat substrates and alternative respiration

As with most bacteria, *E. coli* has the ability to adapt to a number of different environments. Its natural environment is the mammalian intestine where it assists waste processing and the production of vitamin K. It contains several anaerobic respiratory

chains allowing the usage of different final electron acceptors. As described above, each respiratory chain has common components: a dehydrogenase or oxidase that allows the transfer of electrons from a substrate to the quinone pool, and a reductase that accepts electrons from the quinone pool and transfers them to a terminal electron acceptor. *E. coli* is capable of using oxygen as a terminal electron acceptor, but in anoxic environments is can utilize nitrate, nitrite, fumarate, trimethylamine-*N*-oxide (TMAO), or dimethyl sulphoxide (DMSO) as terminal electron acceptors. Most substrates of the Tat pathway are involved in various anaerobic respiratory chains. One of the few Tat substrates for which the structure has been solved is formate dehydrogenase (Jormakka et al. 2002). The molecular dimensions of formate dehydrogenase span up to 8 nm in width and over 15 nm in length (Figure 1.7). The enzyme is believed to form a trimeric unit once it is anchored into the cytoplasmic membrane. It is difficult to imagine how a completely folded protein of this size can cross the cytoplasmic membrane without forcing the the membrane potential to collapse. Since most substrates of the Tat pathway belong to anaerobic respiratory chains and catalyze redox reactions, they typically require a set of complex cofactors and often additional polypeptide subunits. The incorporation of these cofactors often requires specialized chaperones which are only available in the cytoplasm. For example trimethylamine N-oxide reductase (TorA) contains a FeS cluster and a bis-molybdopterin guanine dinucleotide (MGM) cofactor (Mejean et al. 1994). For the assembly process, TorA has its own chaperone, TorD, which greatly facilitates the incorporation of the cofactors and retards the export process (Pommier et al. 1998). TorD binds to the signal peptide and parts of the enzyme.

Since most of the Tat substrates function in anaerobic respiration, it is therefore surprising that the expression of the *tat* operon containing the *tatA,B,C,D* genes is constitutive (Jack et al. 2001). However, there are also several Tat substrates which are involved in other biological processes. As an example AmiA and AmiC are amidases which cleave the peptide moiety of the peptidoglycan layer, a step essential for the completion of cell division. It is the mislocalization of these enzymes which causes the chain-like formation of *tat* mutant cells. The Tat substrate cupper oxidase (CueO) a homologue of laccase confers higher resistance to Cu(II) (Tree et al. 2005) whereas its homologue SufI might contribute to a general robustness of the cell division apparatus in particular under conditions of stress (Samaluru et al. 2007). Further examples of Tat substrates which are not involved in respiration include MdoD, an enzyme that is involved in the synthesis of the osmoregulated periplasmic glucan backbone structures (Lequette et al. 2004) and WcaM, which is involved in lipopolysaccharide biosynthesis. In some pathogenic organisms, Tat substrates are virulence factors (see section on pathogenesis).

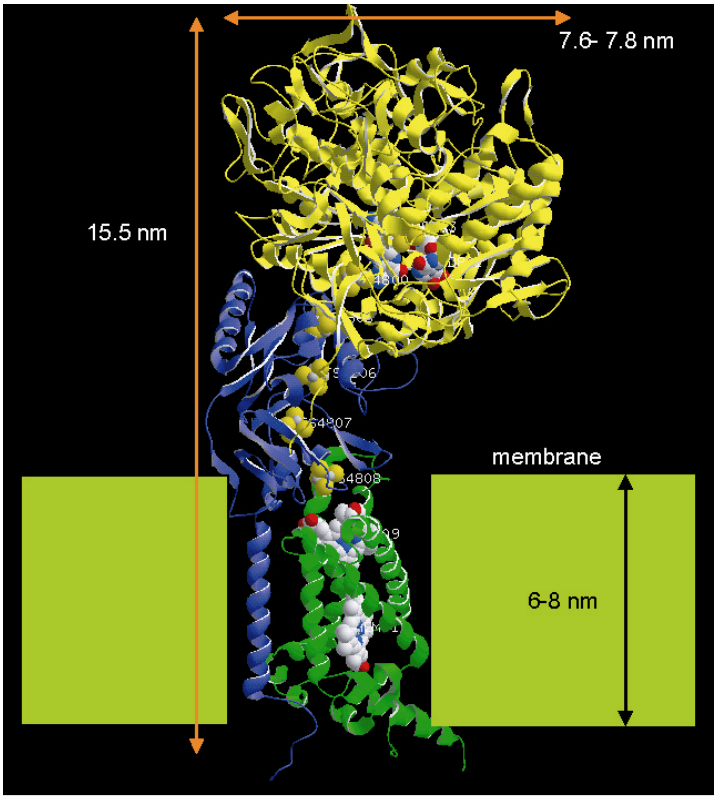


Figure 1.7. Crystal structure one unit of the timeric formate dehydrogenase (1KQF.pdb). Dimensions of this Tat substrate are indicated. FdnG containing the signal peptide is indicated in yellow, FdnH which presumably gets co-transported with FdnG does not contain a signal peptide (blue subunit). The corresponding cytochrome FdnI is colored in green.

Table 1.1. *E.coli* Tat substrates

Protein	***Physiological role***	***Cofactors***	***Co-exported partner***	***Signal chaperone***	***Signal peptide selectivity***
HyaA	Hydrogen oxidation	3×Fe-S clusters	HyaB	HyaE	Tat
HybO	Hydrogen oxidation	3×Fe-S clusters	HybC	HybE	Tat/Sec
HybA	Hydrogen oxidation	4×Fe-S clusters[a]	Unknown	Unknown	Tat/Sec
NapG	Nitrate reduction	4×Fe-S clusters[a]	Unknown	Unknown	Tat
NrfC	Nitrite reduction	4×Fe-S clusters[a]	Unknown	Unknown	Tat
YagT	Unknown	2×Fe-S clusters[a]	YagR[a],	YagQ[a]	Tat

			YagS[a]		
YdhX	Unknown	4×Fe-S clusters[a]	Unknown	Unknown	Tat/Sec
TorA	TMAO reduction	MGD	None	TorD	Tat
TorZ	TMAO reduction	MGD	None	YcdY[a]	Tat/Sec
NapA	Nitrate reduction	MGD, 1×Fe-S cluster	None	NapD[a]	Tat
DmsA	DMSO reduction	MGD, 1×Fe-S cluster	DmsB	DmsD	Tat/Sec
YnfE	DMSO reduction	MGD, 1×Fe-S cluster[a]	YnfG[b]	DmsD[a]	Tat
YnfF	DMSO reduction	MGD, 1×Fe-S cluster[a]	YnfG[b]	DmsD[a]	Tat/Sec
FdnG	Formate oxidation	MGD, 1×Fe-S cluster	FdnH	FdhD[a], FdhE[a]	Tat
FdoG	Formate oxidation	MGD, 1×Fe-S cluster[a]	FdoH[b]	FdhD[a], FdhE[a]	Tat
YedY	TMAO/DMSO reduction	MPT	None	Unknown	Tat/Sec
CueO	Copper homeostasis	4×Cu ions	None	Unknown	Tat/Sec
SufI	Possibly cell division	None	Unknown	Unknown	Tat/Sec
YahJ	Unknown	1×Fe ion[a]	Unknown	Unknown	Tat/Sec
WcaM	Colanic acid biosynthesis	Unknown	Unknown	Unknown	Tat
MdoD	Glucans biosynthesis	Unknown	Unknown	Unknown	Tat/Sec
YcdB	Unknown	Unknown	Unknown	Unknown	Tat/Sec
AmiA	Cell wall amidase	None	Unknown	Unknown	Tat/Sec
AmiC	Cell wall amidase	None	Unknown	Unknown	Tat/Sec
FhuD	Ferrichrome binding	None	Unknown	Unknown	Tat/Sec
YcbK	Unknown	Unknown	Unknown	Unknown	Tat

list of known and predicted *Escherichia coli* Tat substrates (Palmer et al. 2005; Tullman-Ercek et al. 2007), see note

a Predictions inferred by homology, genetic linkage or sequence analysis; (Berks et al. 2003) and (Turner et al. 2004) for reviews.

1.5.3 Signal peptides

Most proteins which are exported across a membrane are synthesized as a precursor form with an N-terminal signal peptide. Signal peptides are evolutionarily well-conserved N-terminal extensions. A signal peptide has three distinct regions (Figure 1.8). The N region harbors a positive charge whereas the hydrophobic H-region comprises the center and the longest part of the signal peptide. The H-region of Tat signal peptides has a less hydrophobic character and is typically longer than that of Sec signal peptides. The N-terminal positively charged region of Tat signal peptides contains the hallmark twin-arginine amino acids with the consensus sequence of S/T-R-R-x-F-x-K. The C-region of both Sec and Tat signal peptides bears the signal peptidase cleavage site which is recognized by the signal peptidase type I (Paetzel et al. 2002). In general, Tat signal peptides typically contain a lysine or arginine residue within this region which has been described as Sec avoidance sequence (Blaudeck et al. 2003). The more positively charged the C-region and beginning part of the mature protein is the higher the chance that the signal peptide will not be targeted to the Sec pathway (Tullman-Ercek et al. 2007).

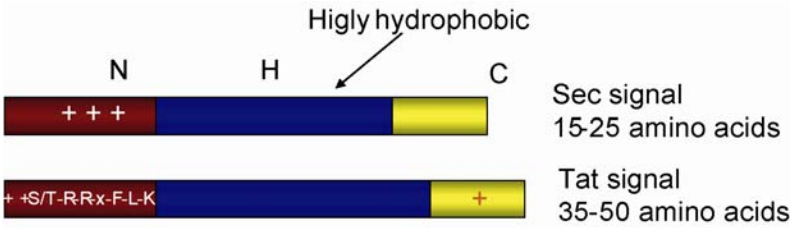


Figure 1.8. Signal peptides. Comparison between a Sec and Tat signal peptides. N, H and C region are indicated. The positive charge within the C region of a Tat signal peptide marks the Sec avoidance signal.

NMR studies on the signal peptide of HiPIP, the high potential iron-sulfur protein (HiPIP) from *Allochromatium vinosum* which had been purified from inclusion bodies and refolded, indicate that the signal peptide is unfolded in solution (Kipping et al. 2003). The purpose of an exposed signal peptide could be to facilitate recognition by the translocation machinery. Different theories about the signal recognition process have been proposed. In one theory, the signal peptide associates with the membrane and scans along the lipid bilayer until it interacts with the receptor complex of the transporter. It is also possible that the signal peptide interacts directly with the receptor complex. Surface plasmon resonance studies of purified precursor proteins (ssTat-GFP) demonstrated electrostatic and hydrophobic interactions of the signal peptide with displayed phospholipids. This evidence implies an initial interaction of the signal peptide with the membrane before the it is recognized by the Tat translocon (Shanmugham et al. 2006).

1.5.3.1 Identification of new Tat substrates based on signal peptide analysis

In a bioinformatics analysis study, the complete *E. coli* genome was searched for proteins containing a Tat consensus sequence within putative signal peptides. This analysis identified 29 putative protein substrates. Signal peptides of all potential substrates were fused to the reporter proteins maltose binding protein (MBP), alkaline phosphatase (AP) and green fluorescence protein (GFP) (see section 1.6.3 under "reporter constructs") and tested for their functionality. Only 11 signal peptides were found to be strictly Tat-dependent whereas the remaining 18 are were shown to be promiscuous, allowing export through either Sec or Tat pathway (Tullman-Ercek et al. 2007). It seems reasonable that the cell throughout evolution maintained flexibility in the translocation of certain enzymes to ensure their correct localization even under conditions in which the Tat export pathway was compromised.

1.5.4 Translocation model

The current model for Tat translocation is based on biochemical data from co-purification, in vitro reconstitution, immunoprecipitation and electron microscopy (Gohlke et al. 2005). In addition, cross-linking studies (Alami et al. 2003) including cysteine cross-linking (Lee et al. 2006; Greene et al. 2007; Punginelli et al. 2007) and genetic analyses (Kreutzenbeck et al. 2007a; Strauch and Georgiou 2007b) have been employed to elucidate the molecular basis of Tat export. In *E. coli*, three proteins TatA, TatB and TatC are essential for the formation of a functional translocon. In resting membranes, TatA, TatB and TatC can be isolated as a 600 kD complex (Bolhuis et al. 2001; de Leeuw et al. 2002). TatA can be pulled out of the membrane as high-molecular weight complexes of varying sizes whereas TatB and TatC tend to co-purify with each other in a stoichiometric manner. Overexpressed, TatB/TatC complexes typically contain variable amounts of TatA oligomers (Bolhuis et al. 2001; de Leeuw et al. 2002). However, more recently it was demonstrated that more moderate expression levels of the translocon components results in two distinct types of complexes in which one comprises only TatB/TatC oligomers and a second complex is established by TatA oligomers (Orriss et al. 2007). Hence, both complexes are stable without each other. This contradicts an earlier theory in which it was postulated that the TatB/TatC subcomplex could only be found in the presence of TatA oligomers due to its instability (Mangels et al. 2005).

The re-adjustment of the composition of the core complexes of the *E. coli* Tat machinery resembles the plant ΔpH pathway. So far, the plant TatA homologue Tha4 was only found associated with the TatB homologue Hcf106 and cpTatC version when the

translocon was active (Cline and Mori 2001). For the plant TatA homologue, it was demonstrated that the aspartate residue in the transmembrane domain is crucial for the recruitment of more TatA oligomers. Cline et al. proposed that substrates first bind to the TatB/TatC receptor complex and are then handed-over to the pore- forming TatA complex. This transfer step is postulated to initiate the final translocation step mediated by TatA oligomers. In the bacterial TatA, there is no aspartate in the N-terminal region, however an aspartate can be found at the N-terminus of TatB. Hence, there might be differences in the mode of TatA recruitment between the thylakoid and the bacterial pathway. The model of a "handing-over" mechanism of the substrate to recruited TatA oligomers might be only partially true since export can still occur when the signal peptide is fused to the C-terminus of the plant TatC homologue (Gerard and Cline 2006). This finding implies that the role of the signal peptide recognition is an independent step from the translocation step. However, the division of the translocation step into two separate events is reflected in the biophysical properties of the translocation: Bageshwar and coworkers demonstrated that translocation requires two electric potentials to take place (Bageshwar and Musser 2007). Therefore it is reasonable to assume that translocation follows a step-wise mechanism (Figure 1.9).

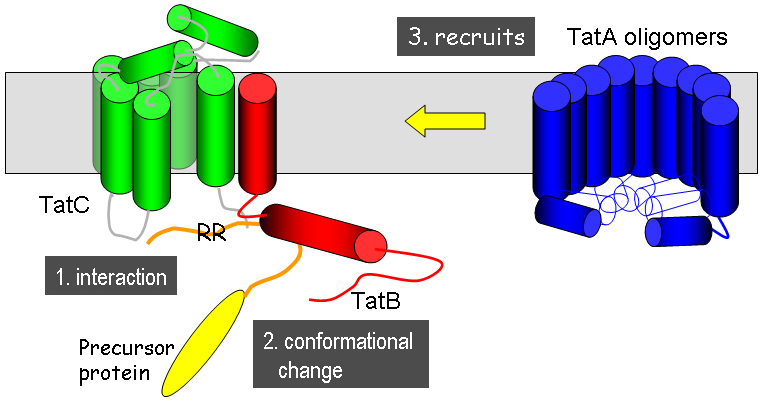


Figure 1.9. Current translocation model. TatB and TatC establish the signal peptide recognition complex. Upon interaction with the signal peptide of the precursor protein, a possible conformational change occurs which signals binding, followed by the recruitment of several TatA oligomers. TatA oligomers assemble in complexes of variable sizes which might depend on the dimensions of the substrate. TatA putatively mediates the actual translocation event.

The role of TatB and TatC functioning as a receptor complex in bacteria was first shown by in vitro cross-linking studies. TatC was found cross-linked to the twin-arginine consensus motif whereas TatB was found to associate with the hydrophobic part of the signal peptide at a later step. Genetic screens by Kreuzenberg (Kreutzenbeck et al. 2007b) and research described in the third chapter of this document confirm the role of TatC as a the site for the initial interaction site of the twin-arginine part of the signal peptide.

1.5.5 Folding quality control

Knowing that the Tat pathway exports folded proteins, provokes the question whether there is a relationship between folding quality and export competence. Does the pathway discriminate between folded and unfolded proteins?

In an elegant study DeLisa and coworkers showed that alkaline phosphatase, in which two intermolecular disulfide bridges must form for it to assume its active dimer conformation, cannot be exported when expressed in cells with a reducing cytoplasm that prevents disulfide bond formation. The same was found to be true for a F_{AB} antibody fragment. A F_{AB} antibody fragment is a heterodimeric protein in which the heavy chain (V_H-C_{HL}) and light chain (V_L-C_L) are linked by an intermolecular disulfide bond; in addition it contains four more intrachain disulfides, two within each of the heavy and

light chains. However, this export deficiency could be eliminated once disulfide bridge formation was allowed in the cytoplasm (DeLisa et al. 2003). The deletions of *gor* and *trxB* in the strain FA113 (MC1000 *phoR* Δ*phoA,Pvu*IIΔ(*malF*)3 *trxB gor552* Tn*10*tet^{r} *ahpC**, F'[*lacI*^{q} *ZYA pro*]) inactivate the thioredoxin and the glutathione reduction pathways which normally maintain the cytoplasm under reducing conditions. Without the reducing power of these two pathways, oxidation can take place in the cytoplasm probably due to Fenton reactions (Bessette et al. 1999). Similarly, Fisher and coworkers showed that maltose binding protein variants (MBP-G32D, MBP-I33P, and MalE31 (G32D/I33P)), which are prone to aggregation of off-pathway folding intermediates, were incompetent for export. The export rate was directly correlated with their folding ability (Fisher et al. 2006). This phenomenon indicates that the pathway functions as a filter for the folding ability of a protein. Based on these observations, Richter and coworkers hypothesized that it might be the exposure of hydrophobic patches which allows the pathway to determine whether a protein is folded or misfolded (Richter et al. 2007). Richter et al. studied the yeast nuclear pore protein Nsp1p which contains a domain with polar FG-repeats which is maintained in a naturally unfolded conformation (Denning et al. 2003). These multiple repeats of 19 amino acids are unstructured and constitute a flexible peptide region which is important for the transport through the nuclear pore. When promoting export of different sized peptide variants of these FG repeats through the Tat pathway, translocation could be demonstrated despite their unfolded nature. However, no export was observed when short hydrophobic stretches were integrated into the FG-repeats. This evidence demonstrates that export competency is not based on the type of fold (unstructured or structured), but is rather simply based on the accessibility of hydrophobic regions on the outside of the protein. Hence, it is conceivable that the

protein quality control of Tat is based on the discrimination of surface-exposed hydrophobic patches, not their folds.

1.5.4 Signal peptide specific chaperones

Several cofactor-containing Tat substrates have their own dedicated chaperone. For example, TorD which assists the maturation of TorA, binds specifically to the TorA signal peptide and also weakly to parts of the mature protein preventing premature export. Its function is to shield the mature TorA protein from recognition by the Tat translocon providing time for the assembly process (Hatzixanthis et al. 2005). TorD binds with specificity to the core region of the TorA signal peptide. The binding step is independent of the presence of the twin-arginine motif, since its affinity for the signal peptide is not reduced by exchanging the twin-arginine for the non-functional twin-lysine motif. Once TorD is bound to the signal peptide, its affinity for GTP increases. However the precise role of GTP in the chaperone function has not been elucidated. It should be noted that there are many protein–protein interactions mediated by GTP binding proteins such as the eukaryotic G protein systems in which guanosine nucleotide exchange factors are required to activate G proteins. Whether there is a similarity between the mode of protein interaction of TorD to the mode of interactions of G proteins has yet to be revealed.

Similarly to TorA, DMSO reductase DmsA and nitrate reductase NarG contain a molybdenum cofactor and an iron-sulfur cluster. Their folding maturation is assisted by the small chaperones DmsD and NarJ, respectively (Oresnik et al. 2001; Chan et al. 2006). Not surprisingly, several of these chaperones are phylogenetically related. Based on phylogenetic analyses, TorD and NarJ have been classified as one group of maturation

chaperones, whereas DmsD and NapD, which assists the folding of the nitrate reductase NapA (Maillard et al. 2007) have been assigned to a second group. At least for NapD, it was recently demonstrated that its molecular role is not only to camouflage the signal peptide; it actively inhibits the transport event before it determines whether the folding maturation is completed (Maillard et al. 2007). As a third group for Tat chaperones, the small chaperones HyaE and HybE were pulled out of protein-protein interaction screens; they assist in the folding maturation of [NiFe] containing hydrogenase 1 (HyaA) and hydrogenase 2 (HybO/HybC), respectively (Dubini and Sargent 2003).

In respect to the proof-reading mechanism of the Tat pathway it is reasonable that this function is divided into two distinct events, one taking place before the interaction with the translocon and performed by the signal peptide-specific chaperone which (a) prevents export before maturation and (b) assists folding, whereas the second part of the quality control is undertaken by the twin-arginine translocon itself, which discriminates against proteins with exposed hydrophobic regions.

1.6 PROTEIN EXPORT REPORTER SYSTEMS FOR MONTORING TAT EXPORT

1.6.1 Disulfide bridge formation

Disulfide bond formation is important for the correct folding and stability of many periplasmic proteins. In *E. coli*, disulfide bridges of proteins are formed in the periplasmic space of the bacterium.

The periplasmic protein DsbA and the integral membrane protein DsbB catalyze the oxidation of cysteine residues in proteins. However, the DsbA/DsbB mediated oxidation acts unspecifically, with DsbA immediately oxidizing consecutive cysteine residues as the protein exits the Sec pore and enters into the periplasmic space (Berkmen

et al. 2005). Proof-reading and correction of disulfide bond formation is performed by the DsbC/DsbD system (Berkmen et al. 2005). DsbC functions as a disulfide isomerase, which re-arranges non-native disulfides to their native configuration. RNase I, MepA (Hiniker and Bardwell 2004), acid phosphatase (AppA) (Berkmen et al. 2005) and the heterologous protein human tissue-type plasminogen activator (tPA) (Qiu et al. 1998) depend on DsbC for folding. DsbC is maintained in a reduced state by its interaction with the integral membrane protein DsbD, which on the other hand receives electrons through its interaction with cytoplasmic thioredoxin. Oxidized thioredoxin is recycled by thioredoxin reductase which uses the NADPH pool as electron source (Collet and Bardwell 2002), see Fig.1.10.

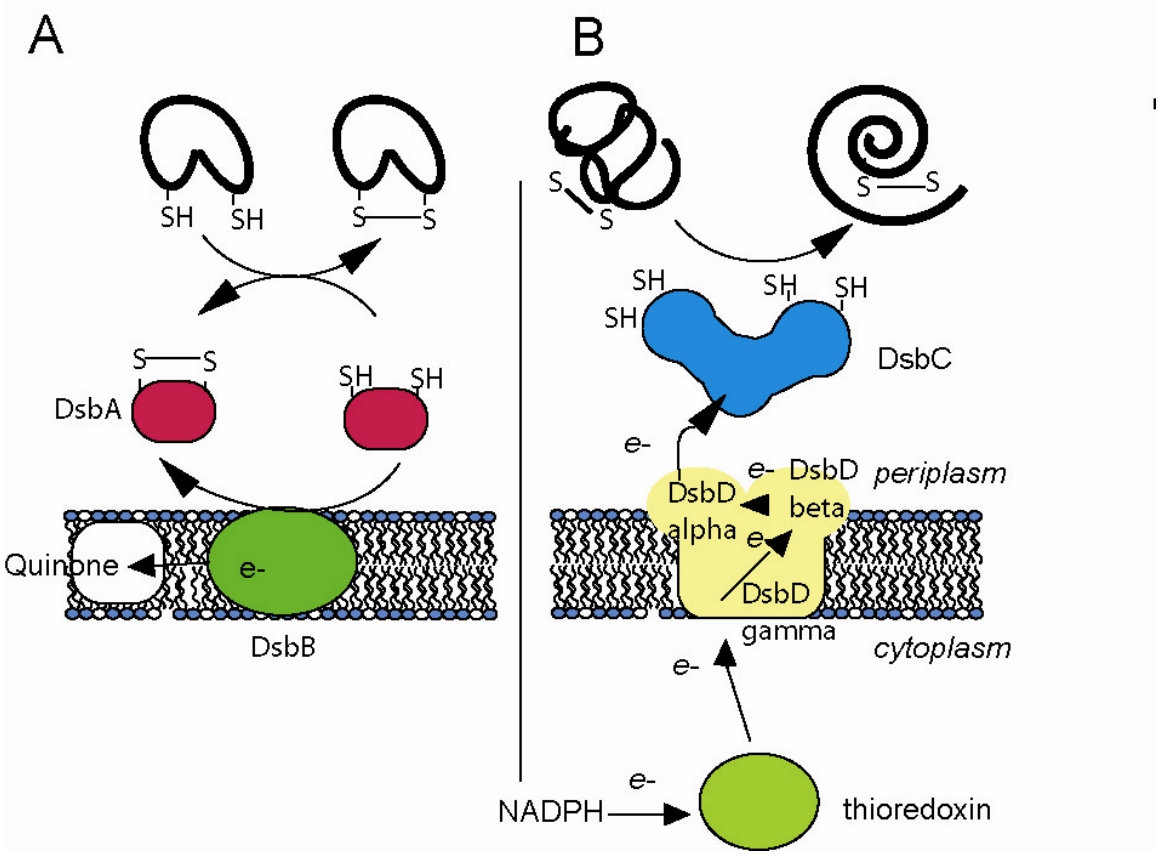


Figure 1.10. Disulfide bridge formation cycle in *E. coli*

1.6.1.1 DsbA

The first step in the formation of disulfide bridges is catalyzed by the 21 kDa enzyme DsbA. Recent studies by Berkman and Beckwith (2005) concluded that this step predominantly takes place while the substrate is being translocated by the Sec-pathway through the cytoplasmic membrane (Berkmen et al. 2005). The active site of DsbA contains a Cys30–Cys33 disulfide bond which has a redox potential of −120 mV and therefore it is a potent oxidant (Rietsch et al. 1997). The formation of a disulfide bridge is initiated by the formation of an unstable mixed disulfide between DsbA and the target protein. The transient complex is then attacked by another thiol-group of the target protein, and reduced DsbA is released. DsbA is subsequently re-oxidized by the action of the inner membrane protein DsbB. DsbB transfers its electrons onto oxidized ubiquinone with oxygen as a final recipient mediated by cytochrome oxidases under aerobic conditions (Figure 1.10). Under anaerobic conditions, DsbB passes electrons to the menaquinone pool in the cytoplasmic membrane which transfers the received electrons onto another reductase e.g. fumarate reductase or nitrate reductase (Collet and Bardwell 2002). Since: (i) the function of DsbA is dependent on its subcellular localization and (ii) its activity can be easily measured by determining the formation of active alkaline phosphatase, DsbA is a useful reporter for detecting the localization of a fusion protein exported either via the Sec (Huber et al. 2005a) or the Tat pathway (see second chapter).

1.6.1.2 Alkaline phosphatase

One of the substrates of DsbA is alkaline phosphatase (AP) encoded by the gene *phoA* (Chang et al. 1986; Akiyama and Ito 1993). AP is a periplasmic enzyme of *E. coli,* exported via the Sec pathway in a post-translational manner. It contains zinc and magnesium as cofactors in its homodimeric form in which each subunit is stabilized by

two intramolecular disulfide bridges. AP cleaves phosphate from a broad range of substrates. The *phoA* gene is part of the *pho* regulon, which is controlled by a two-component regulatory system established by PhoR (sensor kinase) and PhoB (the transcriptional activator protein). At least 30 different genes have been reported to be part of this regulatory network. Under phosphate starvation conditions, the sensor kinase PhoR phosphorylates itself (forming PhoR-P). The phosphate starvation signal is then delivered by the transfer of the phosphate group onto dephosphorylated PhoB, which in turn activates the expression of all genes within the *pho* regulon. Hence, low phosphate concentrations activate transcription from the *phoA* promoter (Martin 2004).

AP is synthesized with an N-terminal 20 amino acid long Sec- signal peptide that is removed during translocation across the cytoplasmic membrane. Following export its free cysteines residues are immediately oxidized by DsbA (Berkmen et al. 2005). AP has been used as a reporter for disulfide bridge formation, for Sec-dependent export, and for periplasmic localization (Manoil et al. 1990). More recently it was used as a reporter for Tat-dependent export in cells with oxidizing cytoplasm (DeLisa, Tullman et al. 2003).

1.6.2 Maltose binding protein

Maltose binding protein belongs to a class of proteins described as periplasmic binding proteins (PBPs) which represents a large family of receptor proteins that recognizes a variety of small molecule ligands within the periplasm. These receptors are involved in carbon catabolism, amino acid and iron transport, chemotaxis and quorum sensing (Tam and Saier 1993). Transport mediated by PBPs typically involves an ABC transport system in which ATP hydrolysis mediates the transport process of the ligand.

MBP, encoded by the gene *malE*, is part of the maltose system that is responsible for the recognition and uptake of α(1→4)-linked glucose polymers (maltodextrins) up to 7 to 8 glucose units (Boos and Shuman 1998). Both substrate-loaded and substrate-free conformation of MBP can bind the integral membrane proteins MalF and MalG which establish the channel. However, an enormous conformational change occurs within MBP upon substrate binding: its two free lobes close up over its substrate upon binding and only this conformational re-arrangement allows MBP to initiate the transport cycle. ATP can bind to two copies of MalK, which is aligned to the cytoplasmic side of the channel. This binding step coincides with a simultaneous opening of both MBP and the periplasmic entrance of the translocation pathway (Austermuhle et al. 2004), channeling maltose inside the cell (Figure 1.11).

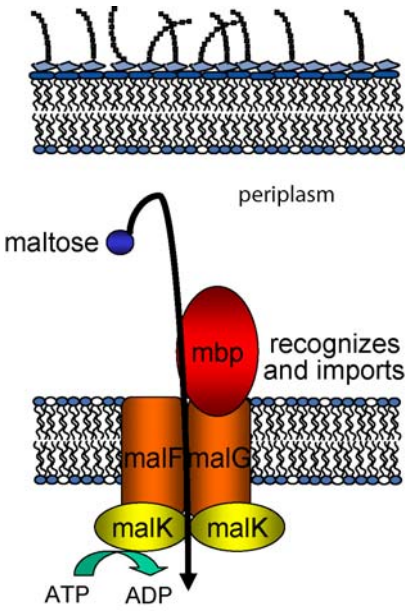


Figure 1.11. Recognition of maltose by the maltose binding protein (MBP)

The requirement for MBP for the catabolism of maltose as a carbon source has been exploited in simple genetic selections for growth on maltose minimal media plates or for screening assays using MacConkey plates. MacConkey plates contain two color indicators which demonstrate whether the cells utilize the carbon source. Genetic screens versus genetic selections (MacConkey vs. minimum media) often have the advantage that the stress-response is not activated which typically reduces the occurrence of spontaneous mutations (Bjedov et al. 2003).

MBP has been extensively used as a reporter for the genetic analysis of the Sec pathway and represents a paradigm of post-translational protein export. MBP folds rapidly and therefore is competent for Tat export. MBP fused to a Tat signal peptide restores growth of *malE* mutants on maltose plates. By deleting one the components of Tat, no growth can be observed indicating that the protein is exported via Tat.

1.6.3 GFP-SsrA

The green fluorescent protein from the jellyfish *Aequorea victoria* has been widely used as a marker for gene expression and protein localization. The chromophore in GFP is generated by a spontaneous cyclization and oxidation of the sequence Ser65-(or Thr65)-Tyr66-Gly67, forming a p-hydroxybenzylidene-imidazolidone (Cody et al. 1993). However, when GFP is exported through the Sec translocon, it assumes a non-fluorescent fold in the periplasm. For unknown reason, fluorescent GFP can only be achieved in *E. coli* when it its folding maturation occurs in the cytoplasm (Feilmeier et al. 2000; Santini et al. 2001). Since Tat substrates fold prior export in the cytoplasmic environment, GFP can be employed as a reporter for Tat export. However, only around 50% of the fluorescent GFP is exported whereas the rest remains in the cytoplasm. For this reason the total cell fluorescence cannot be used to determine the localization of the protein. In

order to detect fluorescence only originating from periplasmic GFP, DeLisa et al. used a C-terminal SsrA degradation tag. The small ssrA RNA (small stable RNA A) was discovered when the expression of murine interleukine-6 (IL-6) in *E. coli* resulted in several truncated versions with each containing an addional C-terminal AANDENYALAA sequence (Tu et al. 1995). Except for the first alanine, all of these amino acids are encoded within the open reading frame of the *ssrA* RNA. It turns out that the bacteria cell uses the aminoacylated-SsrA RNA to exchange mRNA in stalled ribosomes giving rise to a polypeptide with a C-terminal SsrA peptide tag. The 10 amino acids of the SsrA peptide tag directly recruit the tagged truncated protein to the ClpXP degradation machinery (Karzai et al. 2000).

In summary, by artificially tagging GFP with a SsrA sequence, GFP that is not sequestered away from the cytoplasm by transport through the Tat pore is immediately sent to the cellular degradation machinery and eliminated (DeLisa et al. 2002). Thus cell fluorescence is now proportional to the extent of export of GFP-SsrA into the periplasm.

1.7 PHAGE DISPLAY

Phage display has been commonly used as an *in vitro* screening method for isolating polypeptides out of a library of different variants (Smith 1985). Phage display allows the display of polypeptide on the outside of a phage particle and combines the genetic information encoding the displayed protein by placing it inside the phage particle. The displayed protein is typically either encoded in an artificial phagemid or it is directly integrated into the viral genome. A phagemid is simply a plasmid that, in addition to a bacterial plasmid origin of replication, bears a phage-derived origin of replication that promotes packaging into the phage particle. To promote the display of a protein or polypeptide of interest, it is expressed as a fusion to one of the phage envelop proteins.

Since a phagemid does not contain the information for all other phage proteins, cells carrying a phagemid require a superinfection with a helper phage containing remaining phage genes to assure phage propagation (Table 1.2). For phage display, typically filamentous phages such as f1, M13 or fd are utilized. The advantage of using filamentous phage over lytic phage (such as T4, T7 or lambda) is that filamentous phage is directly secreted into the media from which it can be easily purified. The phage assembly process takes place while bacteria are still allowed to grow (around 50% slower duplication time) and phage particles are continuously secreted. Further, the membrane potential is maintained. Secretion into the media simplifies the purification of phage particles tremendously. A polyethylene glycol precipitation step (with high salt concentration) is sufficient to extract the phage from the culture supernatant. This avoids contamination with cytoplasmic proteins during the panning step and removes proteases.

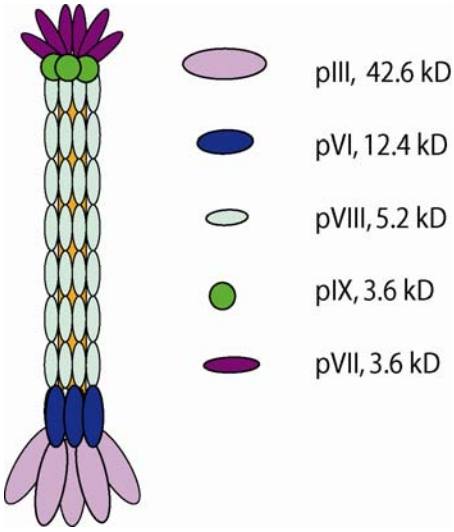


Figure 1.12. Filamentous phage particle

Proteins are typically anchored onto the phage particle by fusion with either the pVIII or the pIII protein. The use of the N-terminus of pIII protein as a fusion partner allows the display of about 3-5 copies of the desired protein per phage particle (Scott and Smith 1990) whereas the fusion to the major coat protein pVIII can result in the display of up to 2700 copies of small (up to six amino acids) peptides (Greenwood et al. 1991) (Figure 1.12). The amount of displayed protein or polypeptide can be varied depending on which phage-display vector is used. However, when using the complete phage genome as a vector and the protein is fused to all pIII or pVIII particles, the viability and infectivity of the phage might be affected depending on the size of the protein.

Table 1.2. Summary of existing phage display vectors

Vector type	***Coat protein used for display***	***Display on all or some copies of coat protein***	***# of coat protein genes***	***Fusion encoded on phage or phagemid genome***
Type 3	pIII	All	1	phage
Type 8	pVIII	All	1	phage
Type 33	pIII	some	2	phage
Type 88	pVIII	some	2	phage
Type 3+3	pIII	some	2	phagemid
Type 8+8	pVIII	some	2	phagemid

During synthesis of a phage particle, the envelop protein pIII, and the assembly proteins pI, pIV and pXI are presumably inserted via Sec into the cytoplasmic membrane,

in a SecA dependent manner (Rapoza and Webster 1993). pIII contains an N-terminal signal peptide. Hence, genes fused to pIII are typically inserted between the signal peptide and the mature part of pIII. pVIII on the other hand inserts spontaneously into the membrane a process that occurs at slower rates depending on the size of the fusion polypeptide (Malik et al. 1998).

Most phage display systems are based on pIII fusions which are exported using the signal peptide of the *Erwinia carotovora* pectate lyase PelB that mediates efficient post-translational Sec transport. However recently it was recently demonstrated that proteins that are highly prone to aggregation, such as Ankyrin Repeat Proteins (DARPins), are more efficiently displayed when targeted to the co-translocational Sec pathway e.g. by a DsbA signal peptide (Steiner et al. 2006).

1.8 BIOTECHNOLOGICAL IMPACT OF THE TAT PATHWAY

The sales of recombinant therapeutic proteins exceed $30 billion per year as of 2004 (Walsh 2005) and are projected to reach $70 billion by the end of the decade. 38% of the newly approved protein-based therapeutics (2003-2006) are engineered. Many therapeutic proteins are produced commercially using *Escherichia coli*. *E coli* is easy, fast and inexpensive to grow in large scale. Its main limitation remains in the lack of a mammalian post-translational modification apparatus especially protein glycosylation which often affects it biological function. Many proteins containing disulfide bonds can be expressed in *E coli* by export into the oxidative periplasm where disulfide bridges are formed. One example of disulfide containing proteins are antibodies.

The special features of the Tat pathway are useful for biotechnological applications related to protein therapeutics. Different technology platforms have been derived using the Tat pathway's inherent folding proof-reading mechanism.

1.8.1 Solubility and folding kinetics

The expression and purification of many eukaryotic proteins is limited by aggregation. Fisher and coworkers developed a screening technology that capitalizes on the folding quality control mechanism of the Tat pathway for the isolation of protein mutants that are resistant to aggregation. Their genetic screen employs a tripartite fusion consisting of the protein of interest, sandwiched between an N-terminal Tat signal peptide and an export reporter such as beta lactamase at its C-terminus. Export of the fusion occurs only when both the desired protein and the reporter are able to fold properly. Using this screening platform, Fisher et al. were able to isolated higher soluble variants of library of randomly mutagenized Aβ42 peptide precursor protein.

In an independent study, Ribnicky and coworkers optimized the folding of a scFv by applying the Tat pathway as a folding filter. A randomly mutaganized library of the scFv 26-10, a scFv antibody fragment binding to digoxin, was screened by flow cytometry for increased expression and binding to fluorescent hapten (Ribnicky et al. 2007). Isolated variants contained mutations only within the framework regions of the antibody fragment and not in the binding loops. In vitro analysis of the folding kinetics revealed that the isolated scFv variant proteins exhibit faster folding kinetics and a lower rate of off-pathway aggregation. Hence, this work demonstrated that the Tat pathway can not only function as a selective filter for higher solubility, but also it can allow the isolation of faster folding protein variants. This is opposite to folding selections

employing export via Sec in which case mutants exhibiting slower folding kinetics are isolated (Huber et al. 2005b).

In summary, export pathways can be exploited in selections for different folding properties in a protein of interest. A genetic screen established on the posttranslational Sec pathway promotes the isolation of slower folding variants, whereas the Tat pathway allows isolation of variants with faster folding kinetics and increased solubility. Both methodologies are of interest for the optimization of the biotechnological expression of therapeutic proteins or to understand the molecular basis of disease-promoting and misfolded proteins as for the case of the amyloid precursor protein involved in the generation of the Alzheimer disease.

1.8.2 Protein-protein interactions

1.8.2.1 A bacterial two-hybrid system

Based on the fact that the Tat pathway is able to export substrates with multiple subunits only one of which contains a signal peptide ("hitchhiker export") (Rodrigue et al. 1999) a new bacterial two-hybrid system was developed (Strauch and Georgiou 2007a). The detection of interacting proteins capitalizes on the folding quality control mechanism of the Twin Arginine Transporter (Tat) pathway and the hitchhiker export of the two interacting proteins. One protein (bait) is expressed as a fusion to a Tat signal peptide whereas the second protein (prey) is fused to a protein reporter that can confer a phenotype only after export into the bacterial periplasmic space. Since the prey-reporter fusion lacks a signal peptide, it can only be exported as a complex with the bait-signal peptide fusion which is capable of targeting the Tat translocon. Using maltose-binding

protein as a reporter, clones expressing interacting proteins can be grown on maltose minimal media or on MacConkey plates. In addition, export of a signal peptide-prey:bait-DsbA complex into the periplasm allows complementation of *dsbA*$^-$ mutants and restores the formation of active alkaline phosphatase, which in turn can be detected by a chromogenic assay. This particular matter will be discussed in chapter 2.

1.8.2.2 Tat-based phage display

Conventional phage display is based on the export of the displayed protein through the Sec pathway which limits the folding environment of the protein of interest (POI) to the periplasmic space. In a Tat-based phage display, the POI is translocated through the Tat pore whereas the pIII protein is sent out through the Sec pathway. Association of the POI is mediated in the periplasm by the interaction of Fos and Jun containing a leucine zipper. The interaction of their adaptor proteins can be stabilized by cysteine residues (Paschke 2005, Tullman-Ercek unpublished results). This system will be discussed in chapter 4.

1.9 THE INVOLVEMENT OF TAT PATHWAY IN PATHOGENICITY

Bacterial pathogenicity depends on the ability of the organism to secret virulence factors. Virulence factors are either displayed on the surface of the cell or secreted into extracellular space. The Tat pathway has been proven to contribute to bacterial pathogenicity to plants and mammals. Homologues of the Tat components have been discovered in *Pseudomonas aeruginosa* (Ochsner et al. 2002), *Listeria monocytogenes* (Desvaux and Hebraud 2006), *Legionella pneumophila* (Rossier and Cianciotto 2005),

Streptomyces coelicolor (Widdick et al. 2006), *Staphylococcus aureus* (Yamada et al. 2007), *Yersinia* (Lavander et al. 2007) and enterohemorrhagic *Escherichia coli* O:157:H7 (Pradel et al. 2003) with varying impact on their virulence.

For example in *Pseudomonas aeruginosa*, 18 different Tat substrates were identified (Ochsner et al. 2002). Tat export was shown to be essential for phospholipases, and proteins involved in pyoverdine-mediated iron-uptake, anaerobic respiration, osmotic stress defense, motility, and biofilm formation. Infection of rat lungs with a Tat deletion mutant of *P. aeruginosa* resulted in low inflammatory responses and no pulmonary abscesses (Ochsner et al. 2002). In the enterohemorrhagic *E. coli* O 157:H7 is was demonstrated that the secretion of the Shiga toxin 1 (Stx1) as well as the H7 flagellin was dependent on a functional Tat pathway (Pradel et al. 2003).

The phytotoxicity of *P. syringae* is mediated by the Tat-dependent export of two extracellular phospholipase C (PLC) that caused major damage to *Arabidopsis thaliana* (Bronstein et al. 2005). *Pseudomonas syringae* pv. tomato DC3000 (DC3000) causes diseases in *Arabidopsis thaliana* and tomato plants, as well as a hypersensitive response in non-host plants such as *Nicotiana tabacum* and *Nicotiana benthamiana*.

Since the Tat pathway is highly conserved among bacteria and archaea and all Tat homologues have been extinguished in the mitochondrial genome of mammalian cells throughout evolution, it could represent a potential target for the development of new therapeutics.

1.9.1 Quorum sensing and the Tat pathway

The Gram-negative bacterium *Providencia stuartii* is responsible for nosocomial and opportunistic infection in humans and causes urinary tract infections. Accumulation of extracellular factors represses the activity of the enzyme 2-*N*-acetyltransferase [AAC(2')-Ia] (Rather et al. 1997). AAC(2')-Ia normally involved in both peptidoglycan and aminoglycoside acetylation as a general house-keeper enzyme. Originally, it was identified as an enzyme which confers the ability of aminoglycoside resistance when overexpressed (Clarke 1993) The precise extracellular compound causing a quorum sensing response has not been identified so far. The regulation of AAC(2')-Ia is influenced by at least 5 different genes which act as negative regulators, one of which is the rhomboid protease AarA (Macinga and Rather 1996). Mutations in the *aarA* gene results in loss of pigment production and in cell division causing chain formation (Rather and Orosz 1994) similar to a *tat* mutant. AarA is a member of the rhomboid family of intramembrane serine proteases which are ubiquitously found in prokaryotes and eukaryotes and are highly conserved. In general, rhomboid proteases are involved in the regulation of growth factor signaling, mitochondrial fusion, and parasite invasion. The only identified substrate of AarA in *P. stuartii* so far is a TatA homologue which has an additional 7 amino acid residues at its N-terminus. Without cleavage of these residues by AarA, the Tat pathway remains non-functional (Stevenson et al. 2007)/ Hence through the activation of the quorum sensing regulated rhomboid protease the activation of the Tat pathway is directly linked to extracellular stimuli in this pathogen.

1.10 RESEARCH OUTLINE

Protein export is employed in a variety of biotechnological applications (such as phage display, cell surface display, antibody expression and folding maturation) which mainly rely on the export via the Sec pathway. Proteins exported by the Sec pathway are extruded across the lipid bilayer in an unfolded state, which excludes the utilization of the cytoplasmic folding machinery. Transport via the Tat pathway can circumvent some of these limitations, because it (i) allows proteins to fold within the cytoplasm and further (ii) the because of the folding quality by the Tat machinery.

As described in chapter 2, we developed a new methodology that capitalizes on the inherent folding proof-reading mechanisms of the Tat pathway and its ability to export more than one subunit via the hitchhiker mechanism. In the new bacterial two-hybrid system, the "signal" (growth or enzymatic activity) can only be detected when the bait-reporter fusion binds to the bait fused to a Tat signal peptide and then both fusion constructs are then exported together via the Tat mediated hitchhiker mechanism. In chapter 4, we describe efforts to use a Tat-based phage display system to isolate GFP variants that bind to different target proteins of interest from libraries of random peptides inserted within a permissive loop of the protein. GFP is a protein that requires folding in the cytoplasm of *Escherichia coli* in order to be functional and therefore it cannot be fluorescently displayed on a phage particle by the conventional phage display system.

The second aspect of this research was aimed at understanding the molecular mechanism of the Tat pathway. Biochemical evidence had suggested that the recognition of the Tat signal peptide is mediated by the receptor complex formed by TatB and TatC. We established genetically the importance of TatC as the crucial component that recognizes the twin-arginine motif within the Tat signal peptide. This work is described

in chapter 3. We identified several amino acid substitutions within TatC that allow an alteration of the substrate specificity of TatC enabling the export of proteins with defective signal peptides. Finally, in the appendix, we describe analysis of TatA as a GFP fusion, combined with proteinase K accessibility analyses, as well as the incorporation of the chemically reactive, non-canonical amino acid azidohomoalanine.

Chapter 2

A bacterial two-hybrid system based on the twin-arginine transporter pathway of *E.coli*

2.1 INTRODUCTION

In recent years, the numbers of genes identified in a broad spectrum of organisms has grown exponentially. A major challenge will be categorizing the corresponding proteins into their functional units within the cell. Several approaches have been used to assign newly identified proteins into cellular networks (Galperin and Koonin 2000; Bork et al. 2004; de Lichtenberg et al. 2005). *In vitro*, interacting proteins can be detected by mass spectrometry or by chromatographic techniques, typically following genetic fusion with an appropriate affinity tag (Rigaut et al. 1999; Gavin et al. 2002; Butland et al. 2005). *In vivo*, genetic techniques for the identification of interacting proteins rely mainly on two-hybrid systems and protein complementation assays. A number of such techniques have been developed and used extensively in *E. coli* (Hu 2001; Karimova et al. 2002; Karimova et al. 2005). So far, the detection of protein interactions in bacteria has capitalized on fusions to transcriptional repressors such as λcI, LexA or AraC, transcriptional activators (involving the recruitment of RNA polymerase or the dimerization of the *Vibrio cholera* ToxR), complementation of biosynthetic enzymes such as dihydrofolate reductase, or signaling enzymes, e.g. the *Bordetella pertussis* adenylate cyclase. Ultimately, the binding of bait and prey is manifested as a change in

the activity of a reporter enzyme which in turn allows for colony formation on selective plates or can be measured quantitatively using chromogenic substrates. A common complication with genetic methods for the discovery of interacting proteins is a high rate of false positives (Fields 2005). False positives can result when one of the two partner proteins alone can trigger a change in the activity of the reporter, for example through spurious binding to DNA that alters transcriptional activation/repression, or as a result of non-specific binding events. Misfolding of the prey or bait gives rise to "sticky" proteins having exposed hydrophobic regions that can engage in non-specific interactions and therefore result in false positives.

This chapter describes the development of a bacterial two-hybrid system that capitalizes on the biological folding quality control mechanism that is inherent to the Twin Arginine Translocation (Tat) pathway for protein export. The Tat pathway of bacteria, archaea and plants is responsible for the translocation of folded proteins across energy transducing membranes (Berks 1996; Palmer et al. 2004; Lee et al. 2006; Sargent et al. 2006). In that regard, the Tat pathway is fundamentally different from the general secretory (Sec) pathway in which proteins traverse membranes in an unfolded conformation by threading through the SecYEG (Sec61αβγ in eukaryotes) pore. The Tat pathway mediates the export of proteins that must assemble with cofactors within the reducing environment of the cytoplasm. In addition, certain Tat precursors associate with other polypeptides in the cytoplasm that do not contain signal peptides. In a process called "hitchhiker export", the two polypeptides form a complex that is targeted to the Tat translocation apparatus and then is exported to the periplasm by virtue of the signal peptide in the precursor (Rodrigue et al. 1999). In the absence of the assembly partner or

when cofactor assembly has been impaired by mutation, protein export through Tat does not take place (Halbig et al. 1999; Rodrigue et al. 1999). These observations together with *in vitro* and genetic evidence point to a protein folding quality control mechanism that is intrinsic to Tat translocation (DeLisa et al. 2003). As a result of the folding quality control feature of the Tat pathway only folded proteins are accepted for translocation across the membrane. This feature of the Tat pathway was recently exploited to develop a screen for protein solubility (Fisher et al. 2006).

We reasoned that the ability of the Tat pathway to discriminate against misfolded polypeptides may be useful for protein interaction analysis. Thus, we developed a two-hybrid system that capitalizes on "hitchhiker export" as well as on the folding quality control features of Tat (Fig. 2.4). We have used two reporter systems for the detection of protein:protein interactions via the "Tat two-hybrid assay": one based on growth on selective media or MacConkey plates and a second, based on an enzymatic assay that relies on the oxidation of alkaline phosphatase in the periplasm, resulting in the formation of active enzyme that is then detected using a chromogenic substrate.

2.2 MATERIAL AND METHODS

2.2.1 Bacterial strains

Bacterial cultures were routinely grown on Luria Bertani (LB) Medium, in liquid culture overnight at 37°C and 250 rpm, and on solid LB medium, containing 1.5% agar overnight at 37°C. Utilized strains are summarized in Table 2.1

Table 2.1. Bacterial strains

Strains	***Genotype***	***Source***
CC160	F^- *thr leu thi lacY galK galT ara fhuA tsx dam dcm supE44*	Lab collection
HS3018	F^-, *araD139* Δ (*argF-lac*) *U169 flbB5301 deoC1 ptsF25 relA1 rbsR22 rpsL150 thiA malT(Con)-1* Δ*malE444*	Shuman 1982
MC1000	*araD139* Δ*(araA-leu)7679 (codB-lac)X74 galE15 galK16 rpsL150 relA1 thi* [F^+ *proAB, lacIq, lacZ ΔM15, Tn(*tetR*)*]	Lab collection
MCA	MC1000, Δ*dsbA::kan*	Lab collection
Xl-1 blue	*recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac,* [F^+ *proAB, lacIq, lacZ ΔM15, Tn(*tetR*)*]	Stratagene

2.2.2 Plasmid constructions

Plasmids used in this work are summarized in Table 2.2 and oligonucleotides in Table 2.3.

Table 2.2. Plasmid constructs

Plasmids	***Genotype***	***Reference***
pEMSI	Cm, pACYC replicon, *trc* promoter	This study
pEMS11	Cm, pACYC replicon, p_{trc} wb*fos::*Δ*ssmalE*	This study
pEMS12	Amp, pTrc99 replicon, , p_{trc} *sstorA::*wb*jun*	This study
pEMS13	Cm, pBAD33 replicon, p_{BAD} wb*fos::*Δ*ssdsbA*	This study
pEMS14	Cm, pBAD33 replicon, p_{BAD} Δ*ssdsbA::im2*	This study
pEMS15	Amp, pTrc99 replicon, p_{trc} *sstorA::E2*	This study
pEMS16	Amp, pTrc99 replicon, p_{trc} *sstorA(KK)::*wb*jun*	This study

Table 2.3. Oligonucleotides

Name	***Sequence***
EMS11	AGCTGGACGTCCGTACGTTCGAACCGACAAAAC
EMS12	GCGGCGATCGATCGACTGCACGGTGCACCAATG
EMS13	GCGTAGTCATGAACAATAACGATCTCTTTCAGG
EMS14	GCGCGCTCTAGAAGCGTCAGTCGCCGCTTGCGCCGC
EMS15	GCGCGTCATGAACAATAACGATCTCTTTCAGGCATCAAAGAAACGTTTTCTGGC ACAAC
EMS16	GCGATGTCTAGAATCGCCCGGCTAGAGGAAAAA
EMS17	CGATGAAGCTTTCAGTGGTTCATGACTTTCTGTTT
EMS18	GCGGCGTCTAGAAAAATCGAAGAAGGTAAACTGGTA
EMS19	GCGGCGAAGCTTTTACTTGGTGATACGAGTCTGCGC
EMS110	GCGGCGGAGCTCATGCTGACCGACACCCTGCAGGCGG
EMS111	GCGATGTCTAGAGTGTGCCGCCAGGATGAA
EMS112	GCGATGGAATTCGAGCTCTTAAAGAGGAGAAAG GTAC
EMS113	GCGCGCGTCGACATGGCGCAGTATGAAGATGGTAAACAG
EMS114	GCGGCGAAGCTTTTATTTTTTCTCGGACAGATATTTC
EMS115	GCGGCGTCTAGAGAGAGTAAACGGAATAAGCCAGGGA
EMS116	GCCCGCAAGCTTTTATTTTTCAAACTGCGGATGGCTCCACTTACCCCGATGAAT ATCAATCGCTCGCTTAGGTGTGGTC
EMS117	GCGCGGGAGCTCAGGAGGAATTCACCATGGCGCAGTATGAAGATGGTAAACAG T
EMS118	GCGCGCTCTAGAGGATCCTTTTTTCTCGGACAGATATTTCACTGTATCAG
EMS119	GCGGCGTCTAGAATGGAACTGAAACATAGTATTAGTGATTATACCG

Plasmid pEMSI was constructed by removing the *araC* regulator gene, the p_{BAD} operator and promoter region of pBAD33 (Guzman et al. 1995) and replacing them with a 220 bp fragment 5' of the ATG codon in pTrc99A. The latter fragment (with *ClaI* and *HindIII* restriction sites) was amplified with the primers EMS11 and EMS12. wb*fos* was amplified with EMS110 and EMS111, and Δ*ssmalE* with EMS18 and EMS19. wb*fos* was inserted upstream of Δ*ssmalE* into pEMSI using the *Sac*I, *Xba*I and *Hind*III cut sites for these two genes, resulting in pEMS11 (Figure 2.1).

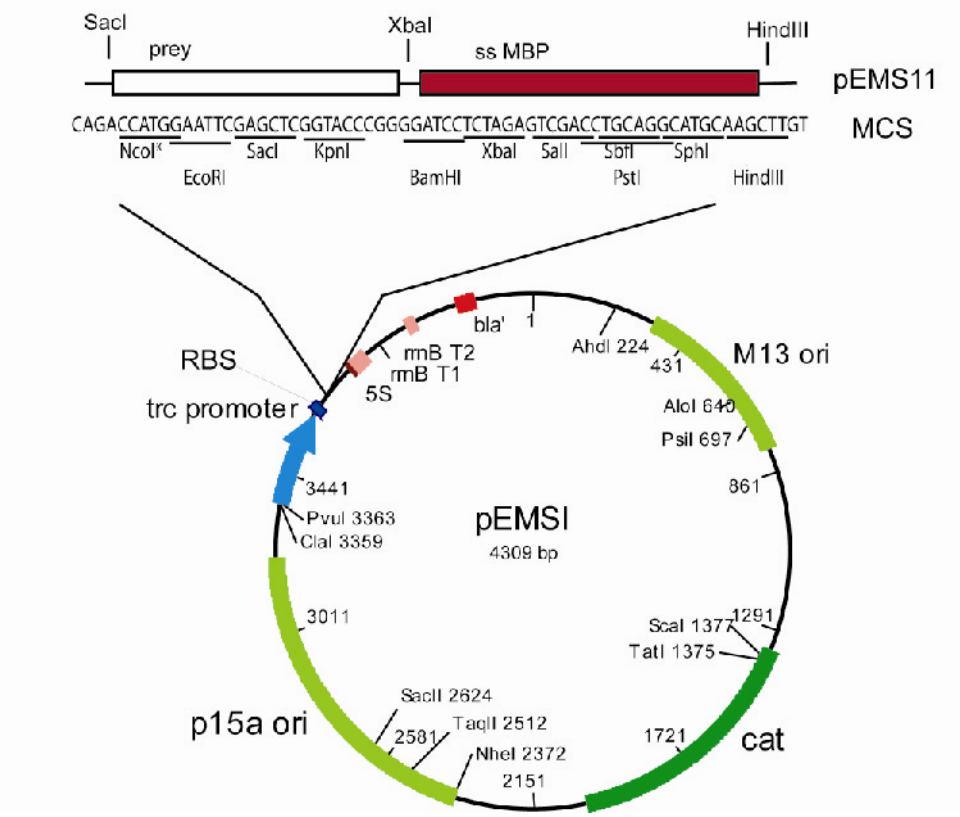


Figure 2.1. Plasmid map of pEMSI

The intermediate plasmid construct pTrc99-ssTorA was generated by amplifying *sstorA* with the primers EMS13 and EMS14 and inserting into pTrc99 between the *Nco*I and *Xba*I sites. wb*jun* was amplified with primers EMS16 and EMS17 and ligated into pTrc99-ssTorA using the *Xba*I and *Hind*III sites, giving rise to pEMS12. The KK version of the TorA signal peptide was generated as above, but using EMS15 as 5' primer, generating plasmid pEMS16.

For the construction of the wbFos-ΔssDsbA fusions, wb*fos* was amplified with the primers EMS112 and EMS111, digested with *Sac*I and *Xba*I and inserted into pBAD33. In addition, Δ*ssdsbA* was cloned directly downstream using the primers EMS113 and

EMS114, producing pEMS13. The DNase domain of colicin E2 was amplified from the plasmid ColE2-P9 (Mock and Schwartz, 1978) with the primers EMS115 and EMS116 containing a codon change of H575A and a 5' strep2 tag sequence. After digestion with *Xba*I and *Hind*III, the DNase domain was inserted downstream of Δ*sstorA* in pTrc99-ssTorA, generating pEMS15.

Plasmid pEMS14 was generated by PCR amplification of Δ*ssdsbA* with the primers EMS117 and EMS118 using genomic DNA and then insertion of Δ*ssdsbA* in between the *Sac*I and *Xba*I site of pBAD33, whereas immunity protein 2 was inserted downstream of Δ*ssdsbA* between *Xba*I and *Hind*III using primers EMS119 and EMS120 containing an additional FLAG tag at the 3' site. The 5' position of Δ*ssdsbA* allows an easy exchange of bait and prey since the restriction sites are compatible with the signal peptide construct.

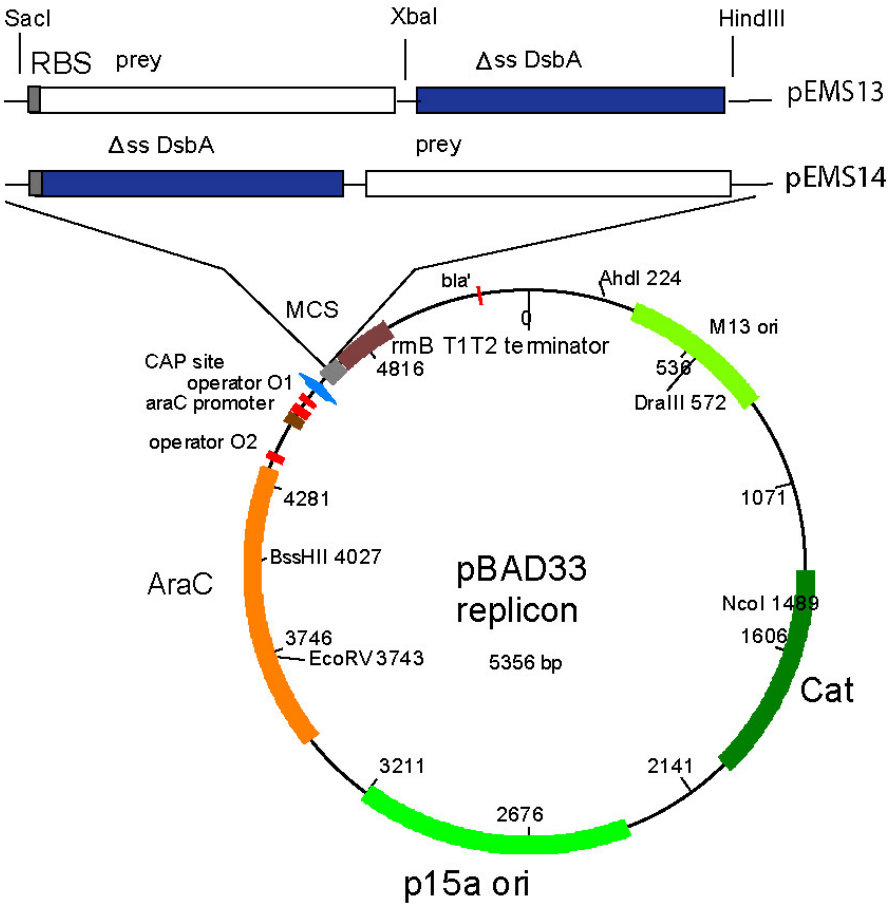


Figure 2.2. Plasmid map of pBAD33 constructs containing the DsbA reporter fusions

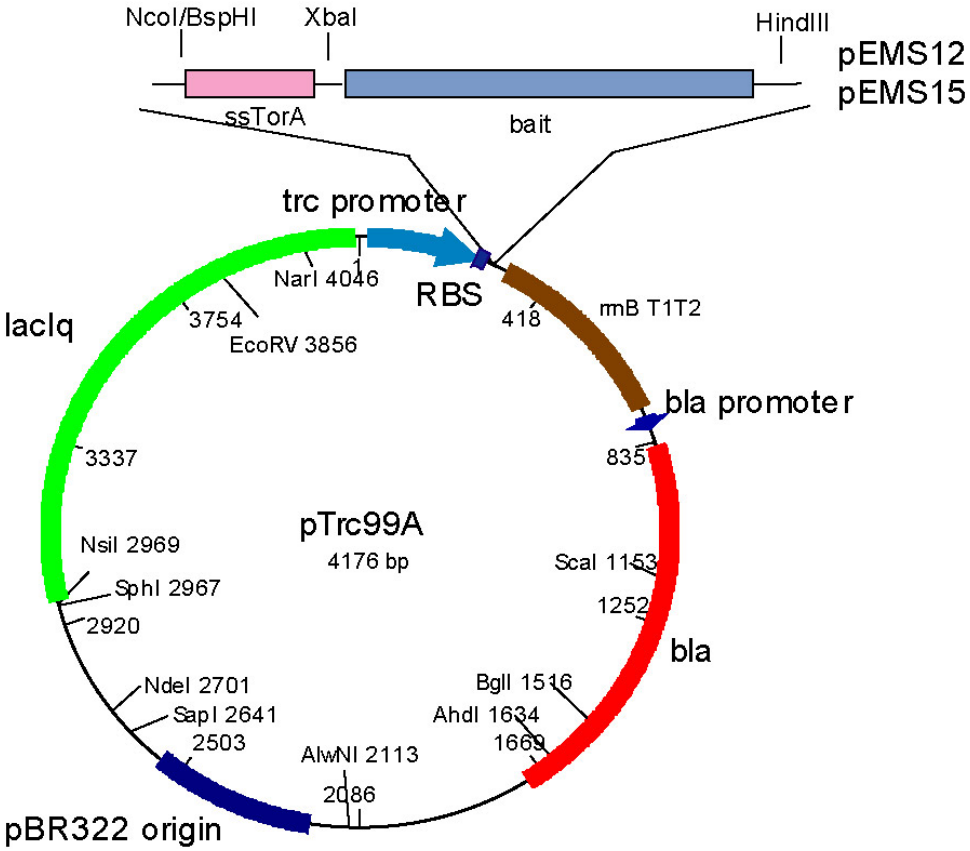


Figure 2.3 Plasmid map of the bait constructs

2.2.3 Media and buffer composition

Table 2.3 Summary of growth media

Medium	Component	% w/v or %v/v or mM
Luria Bertani medium (LB) (pH7.0)	tryptone	1.0%
	yeast extract	0.5%
	NaCl	1.0%
M9 medium	NH_4Cl	0.1%
	NaCl	0.05%
	KH_2PO_4	0.3%
	Na_2HPO_4 (anhydrous)	0.6%
YENB medium	yeast extract	0.75%
	nutrient broth	0.8%
MacConkey medium	Peptone	1.7 %
	Proteose Peptone	0.3 %
	Bile Salts No. 3	0.15 %
	NaCl	0.5 %
	Agar	0.135 %
	Neutral Red	0.003 %
	Crystal Violet	0.0001 %
	pH 7.1	
Micronutrients	$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	0.0037 %
	$CoCl_2$	0.0084 %
	$MnCl_2 \cdot 4H_2O$	0.0158 %
	H_3BO_3	0.0247 %
	$CuSO_4$	0.0025 %
	$ZnSO_4$	0.0018 %
MOPS medium (induction medium for alkaline phosphatase)	10x MOPS salts	1x
	glucose	0.2% (w/v)
	casamino acids	0.2% (w/v)
	adenine	20 µg/ml
	thiamine	0.5 µg/ml
	M9 neutral phosphate buffer	0.1 mM
MOPS salts (10x)	3-(*N*-morpholino)propanesulfonic acid (pH 7.4) (MOPS)	400 mM
	Tricine (pH 7.4)	
	$FeSO_4 \cdot 7H_2O$	40 mM
	NH_4Cl	0.1 mM
	K_2SO_4	95 mM
	$CaCl_2 \cdot 2H_2O$	2.8 mM
	$MgCl_2 \cdot 6H_2O$	5 µM
	NaCl	5.3 mM
		0.5 M

2.2.4 Preparation of electrocompetent cells

Flasks containing 500 ml of YENB low salt medium were incubated with 3 ml of an overnight culture of the desired strain. Cultures were grown at 37˚C or 30˚C as described, until an O.D.600 of 0.6 was reached. Cells were harvested by centrifugation at 10,000g for 15 min at 4°C. Cells were re-suspended in ice-cold dH_2O and centrifuged again. This washing step was performed once more and a third time with 15 % glycerol instead of dH_2O. The final centrifugation step was extended to 20 min and cells were re-suspended into a final volume of 1.125 ml per 500 ml original culture volume in 15% glycerol. The cell suspension was aliquoted and immediately stored at -80°C.

2.2.5 Alkaline phosphate activity assays

Cells were grown overnight in low phosphate MOPS minimum media (Sambrook et al. 2000) with Peptone P (USBiological) as an amino acid source, subcultured (1/20) into 1 ml low phosphate medium and induced with 0.04% arabinose at OD_{600} 0.4. Around 6 h after induction, cultures were diluted with media to OD_{600} 1.0; 20 µl of cell cultures were transferred to a 96 well plate and mixed with 30 µl of lysis buffer, a 2:1 mixture of B-PER (78248 Pierce Biotechnology, Inc.) with 0.4 M iodoacetamide (Sigma) for 30 min. As substrate for AP, 200 µl of 250 µg/ml p-nitrophenyl phosphate (Sigma) in 0.250 M Tris-HCl pH 8 was added to each well and hydrolysis was measured at A_{405} on a plate reader (BioTek).

2.2.6 Carbon source selection assays

Cells containing the desired plasmid combinations were grown overnight at 37˚C in LB medium, washed in M9 minimum media, diluted 10^6 fold and plated.

For minimal media growth selections and plating on MacConkey agar, the growth media was supplemented wth 0.4% maltose. On minimal media plates with maltose as the carbon source, colonies appeared after 4-5 days of incubation at 37˚C. For MacConkey agar based plates, the results can be seen after 18 h.

2.3 RESULTS

As indicated in Figure 2.4, the bait is fused in frame to the C-terminus of a Tat signal peptide. One of the best studied Tat signal peptides is ssTorA of the *E. coli* trimethylamine N-oxide reductase (TorA) and it is selected for the present study. The prey was expressed as a fusion to a reporter that has to be translocated to the periplasm for function. Since the prey-reporter fusion does not contain a signal peptide, it is unable to interact with the Tat translocation pore and therefore remains sequestered within the cytoplasm. However, binding of the prey to the signal peptide-bait fusion results in the formation of a protein complex that can be targeted to the Tat translocon and then be exported into the periplasm.

The maltose-binding protein, MBP, is encoded by the *malE* gene and has been used extensively as a fusion partner in protein export studies and to aid protein solubility (Boos and Shuman 1998; Kapust and Waugh 1999; Blaudeck et al. 2003). The export of MBP to the periplasm can be detected easily on MacConkey agar plates or by growth on maltose as the sole carbon source. Signal sequence-less ΔssMBP prey fusions were expressed under the control of the p_{trc} promoter in the low copy number plasmid pEMSI (pACYC origin of replication). The signal peptide-bait fusion was similarly transcribed from a p_{trc} promoter on a compatible plasmid vector containing a pBR322 origin. To avoid cell toxicity due to the high level of expression of the two fusion proteins, we relied on leaky transcription from the p_{trc} promoter in the absence of inducer. As discussed below, the basal level of transcription from uninduced cells is sufficient to give a high signal to noise ratio.

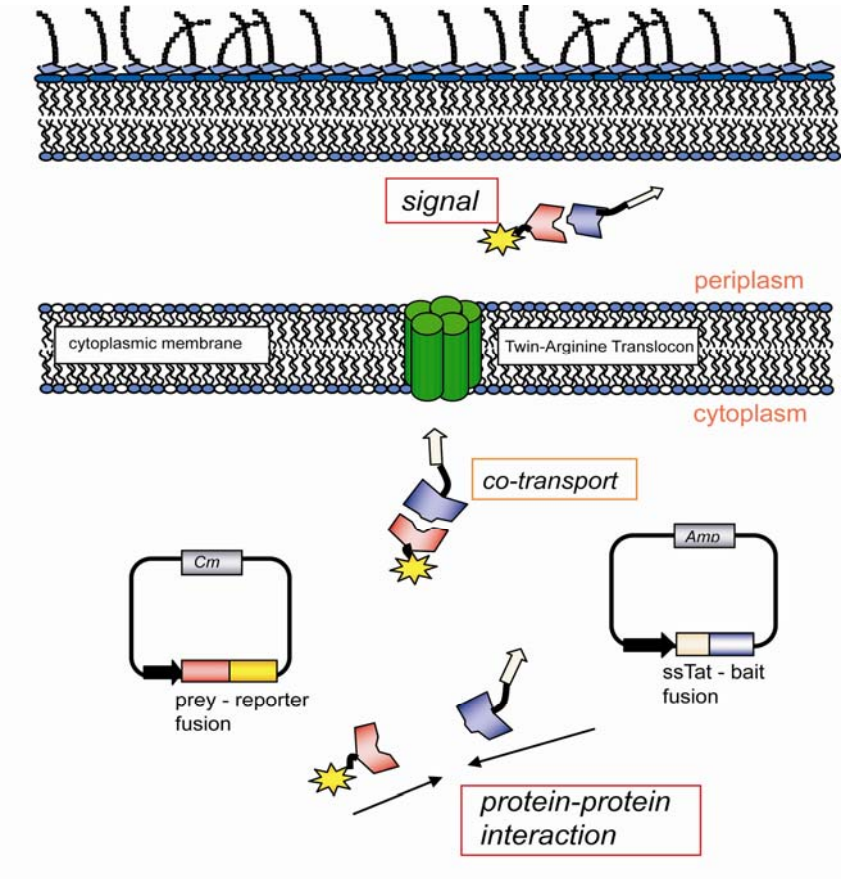


Figure 2.4**.** Generalized scheme for Tat-based two-hybrid system. A pACYC-derivative plasmid expresses the prey fusion to a reporter protein; the fusion can be either C- or N-terminal to the reporter. A compatible pBR322 replicon expresses the bait C-terminal of a Tat signal peptide. Interaction of the bait and prey allows export of the resulting complex through the twin-arginine transporter. In the periplasm, the specific activity of the reporter can be detected.

2.3.1 Maltose binding protein as a reporter for export

As model bait and prey proteins, we used the leucine zipper domains of the transcription factor wb*fos* (rat, aa 161-200) and wb*jun* (rat, aa 227-315) respectively (Abate et al. 1990; Crameri and Suter 1993). These two domains represent one of the prototypical systems for interacting proteins (Kouzarides and Ziff 1988) with a K_D of 70 nM (Oyama et al. 2006). Plasmids encoding wbFos-ΔssMBP (pEMS11) and ssTorA-wbJun (pEMS12) were transformed into the *E. coli malE*$^-$ strain HS3018 and plated on MacConkey agar plates. Bacteria containing either plasmid alone gave rise to yellow-orange colonies revealing their inability to utilize maltose as the carbon source. In contrast, cells co-transformed with pEMS11 and pEMS12 acquired the ability to grow on maltose and formed red colonies after growth for 14 h at 37˚C (Fig. 2.5.A.2). Substitution of the twin arginine dipeptide with a pair of lysines completely abolished Tat export (DeLisa et al. 2002; Ize et al. 2002): cells expressing the ssTorA(KK)-wbJun (pEMS16) gave rise to yellow-orange colonies (Fig. 2.5.A.1) indicating that the *mal*$^+$ phenotype shown in Fig. 2.A.2 was dependent on Tat export. Additionally, only cells transformed with both plasmids and encoding a wt ssTorA signal sequence formed colonies on minimal media plates with 0.4% maltose as the carbon source (Fig. 2.5.A.4). No colony growth was observed for the ssTorA(KK) fusion whereas the wt ssTorA showed growth after 4 days, consistent with the phenotype observed on MacConkey plates (Fig. 2..5.A.3)

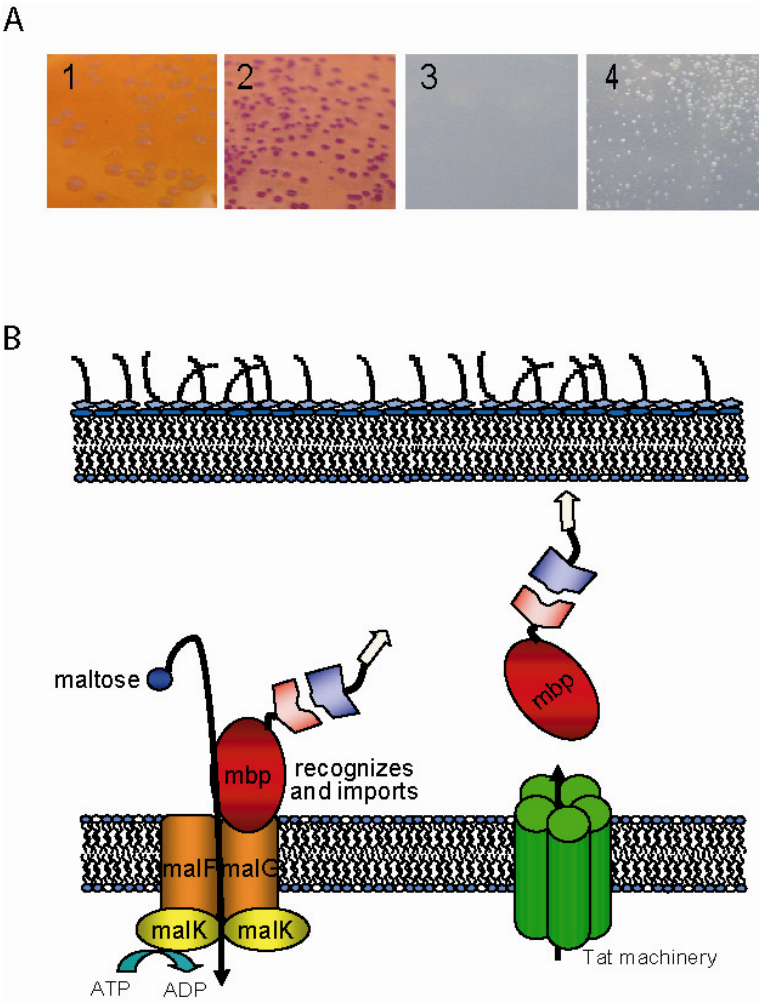


Figure 2.5. Detection of interacting proteins via Tat two-hybrid using MBP as reporter. (A) Use of MalE as a reporter: Panel 1: *E.coli* HS3018 (*malE*) cells co-expressing ssTorA(KK)-wbJun fusion with a non-functional Tat signal peptide (from pEMS16) and wbFos-MBP (from pEMS11) grown on MacConkey plates with 0.4% maltose. Panel 2: *E.coli* HS3018 (*malE*$^-$) co-expressing ssTorA-wbJun (pEMS12) with wbFos-MBP (pEMS11) grown on MacConkey plates with 0.4% maltose. Panel 3: Cells as in Panel 1 plated on M9 media with 0.4% maltose and incubated for 4 days at 37˚C. Panel 4: Cells as in Panel 2 grown on maltose minimal media plates as in panel 3. (B) Interaction of the prey-ΔssMBP fusion with ssTorA-bait results in a complex that is co-translocated into the periplasm by the Tat pathway via the ssTorA signal peptide. In the periplasm MBP channels maltose to the maltose uptake apparatus which allows the cells to catabolize maltose as a carbon source.

2.3.2 Disulfide bridge formation as indicator for protein interactions

We then sought to develop an enzymatic assay for the export of the ssTorA-bait:prey-reporter complex into the periplasm. However, we could not take advantage of the widely used periplasmic reporter alkaline phosphatase (AP) because it has to fold in the cytoplasm before it can be exported via Tat (DeLisa et al. 2003). Cytoplasmic folding of AP can only be accomplished in specialized oxidizing mutant strains (Bessette et al. 1999). In addition, the large size of the folded AP, 94 kDa, (Bradshaw et al. 1981) limits the size of the bait and prey proteins in the fusions because it is unlikely that the Tat pore can accommodate signal peptide-bait:prey- complexes larger than 160-180 kDa (Santini et al. 2001). Similarly, β-lactamase which has been used as a periplasmic reporter enzyme is not suitable for our purposes because: (a) β-lactamase can fold into an active conformation within the cytoplasm and therefore the enzymatic activity in total cell lysates is not indicative of export and (b) a low level export and resistance to ampicilin is observed in cells expressing signal sequence-less β-lactamase (Bowden et al. 1992).

We elected to investigate the utility of the periplasmic cysteine:disulfide oxidoreductase DsbA as a periplasmic reporter (Fig. 2.7.A). In bacteria, the oxidation of protein thiols to form disulfide bonds normally occurs in the periplasm and is catalyzed by DsbA (Kadokura et al. 2003). DsbA, a thioredoxin superfamily protein, is a highly efficient catalyst of disulfide bond formation. Its activity depends on recycling by the membrane-bound DsbB enzyme which in turn transfers electrons to the quinones and ultimately, to the respiratory chain (Nakamoto and Bardwell 2004). The folding of alkaline phosphatase in the periplasm is dependent on oxidation by DsbA. In the *E.coli* *dsbA*$^{-}$ mutant MCA (MC1000 Δ*dsbA::kan*) grown in low phosphate MOPS minimum media, AP activity is about 20-fold lower compared to the isogenic control (Figure 2.C).

The level of residual AP activity is partly dependent on the Fe(II) concentration in the growth media (data not shown).

wbFos-ΔssDsbA was expressed from the arabinose inducible promoter in plasmid pEMS13 to allow differential expression of the bait and prey fusion proteins. Plasmids pEMS12 and pEMS13 were co-transformed into *E.coli* MCA and grown in low phosphate media to induce the expression of AP from the chromosomal *phoA* gene. The synthesis of the wbFos-ΔssDsbA expression from pEMS13 was induced during logarithmic phase ($O.D._{600}$ ~0.4) with 0.04% arabinose. Cells were harvested around 6 hours after induction, lysed, free thiols were blocked with the sulfhydryl alkylating agent iodoacetamide to prevent oxidation of any reduced AP after lysis (Derman and Beckwith 1991) and the enzymatic activity was determined using the chromogenic substrate pNPP (Fig. 2.C). Cells expressing wbFos-ΔssDsbA alone had an AP level slightly above background whereas co-expression of ssTorA-wbJun from pEMS12 resulted in nearly complete restoration of AP activity to the level observed in cells expressing authentic DsbA from the chromosomal gene. Colicin E2 (E2) and the immunity protein 2 (Im2) constitute a pair of toxin:antitoxin proteins that interact with high affinity ($K_D=10^{-15}$ M). Im2 binds to the endonuclase domain of Colicin E2 with a 1:1 stoichiometry to inhibit cleavage of DNA prior to export from the cytoplasm. Garinot-Schneider and coworkers described several mutations that abolish DNase activity (Garinot-Schneider et al. 1996). A gene encoding the 86 aa Im2 protein was fused 3' of Δss*dsbA* gene giving rise to pEMS14. The C-terminal cloning of Im2 introduced the same restriction sites as those used for the signal peptide-bait plasmid; thus, bait and prey genes can be easily interchanged between the two plasmids encoding the signal peptide and the reporter fusion. The H575A mutation that disables the active site was introduced in ColE2 (aa 479-582) and the resulting gene was fused to the C-terminus of ssTorA (pEMS15). Cells

expressing both ssTorA-E2 in pEMS15 and ΔssDsbA-Im2 in pEMS14 exhibited high AP activity (25-fold above background, Figure 2.C). As expected, expression of ΔssDsbA-Im2 alone gave a very low level signal. We further examined whether the expression of non-interacting proteins would give us any AP activity. We tested wbFos-ΔssDsbA with the endonuclease domain of colicin E2, ssTorA-wbJun with immunity protein 2, and additionally examined whether the immunity protein could have any dimerization properties using ssTorA-Im2 and Im2-ΔssDsbA. None of these three protein fusion pairs gave signal above background (Figure 2.6 and data not shown).

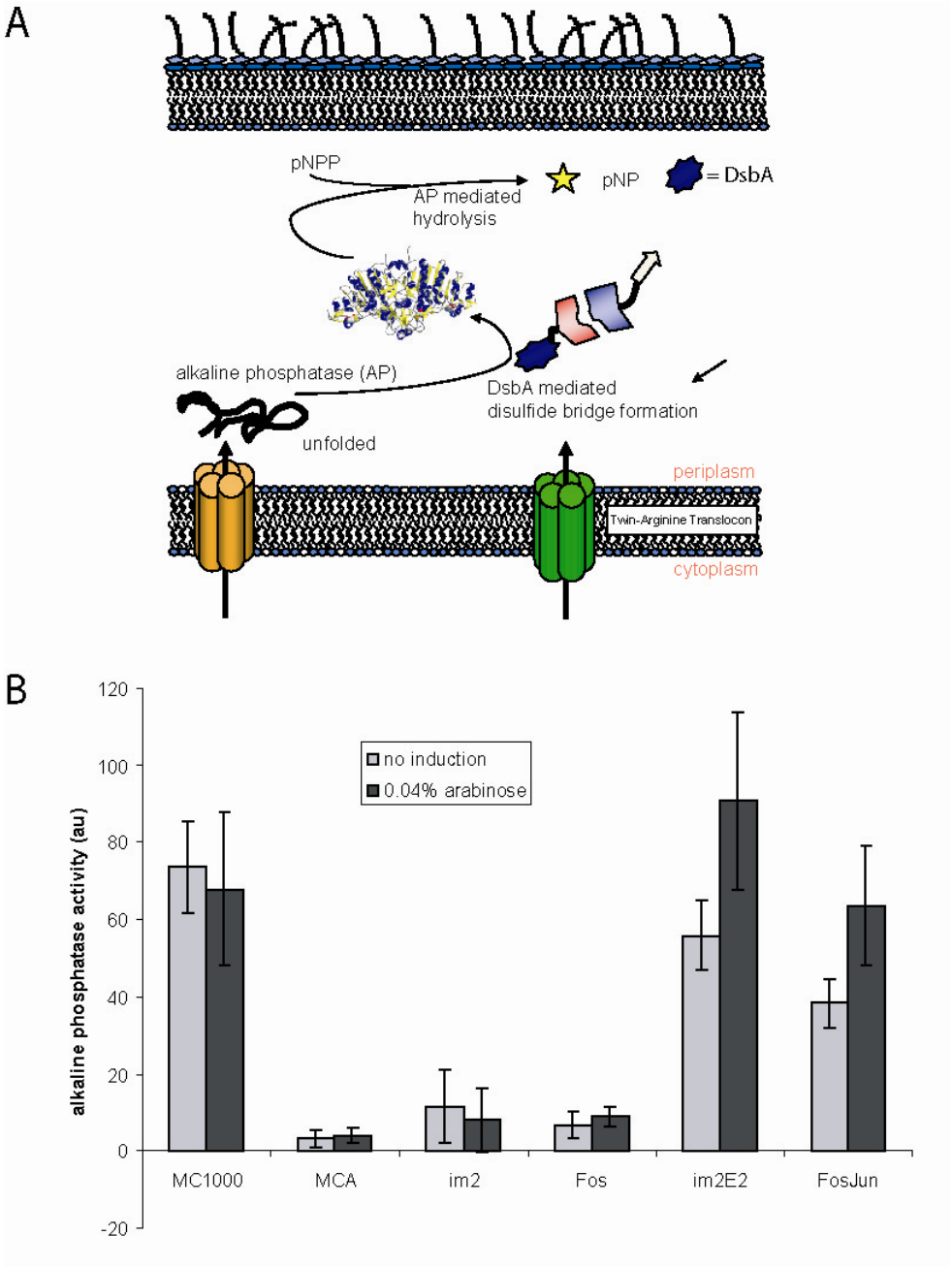

A
pNPP
pNP
= DsbA
AP mediated
hydrolysis
alkaline phosphatase (AP)
DsbA mediated
disulfide bridge formation
unfolded
periplasm
Twin-Arginine Translocon
cytoplasm
B
alkaline phosphatase activity (au)
120
100
80
60
40
20
0
-20
no induction
0.04% arabinose
MC1000
MCA
im2
Fos
im2E2
FosJun

Figure 2.6. Detection of protein:protein interactions by an enzymatic assay using DsbA as a reporter. (A) Interaction of the prey-ΔssDsbA fusion with ssTorA-bait results in the formation of a complex that is translocated into the periplasm by the Tat pathway via the ssTorA signal peptide. In the periplasm DsbA catalyzes the oxidative folding of AP the activity of which can be detected using the chromogenic substrate pNPP. (B) Alkaline phosphatase activity in arbitrary units (au). Cells were grown in MOPS low phosphate media with or without 0.04% arabinose, harvested after 6 hours and the AP activity was measured as described in the Material and Methods section. *E. coli* MC1000 (positive control); MCA (MC1000 *ΔdsbA::kan*) (negative control); Im2: *E.coli* MCA expressing ΔssDsbA-Im2 from pEMS14; Fos: MCA expressing wbFos-ΔssDsbA from pEMS13. Im2:E2: co-expression of the ΔssDsbA-immunity protein 2 fusion from pEMS14 together with the ssTorA-endonuclease domain of colicin E2 (H575A) fusion from pEMS15; Fos:Jun: co-expression of ssTorA-wbJun from pEMS12 with the wbFos-ΔssDsbA fusion (pEMS13); Jun:Im2: co-expression of ssTorA-Jun from pEMS12 with immunity protein2 fused to ΔssDsbA (pEMS14).

2.4 DISCUSSION AND CONCLUSIONS

We have developed a new genetic assay for the detection of protein:protein interactions by capitalizing on the unique features of the Tat export pathway of *E.coli*. To our knowledge, this is the first protein interaction assay that takes advantage of protein translocation across an energy transducing membrane, namely the cytoplasmic membrane of *E.coli.* Extensive studies have shown that only folded proteins can be translocated via the Tat pathway, including proteins comprising of two or more interacting polypeptides only one of which contains a signal sequence. Rodrigue et al. (1999) first coined the term hitchhiker export after observing that the large subunit HybC of the *E. coli* hydrogenase 2, which does not contain a signal peptide, was translocated as a complex with HybO which contains a 37 aa long Tat signal peptide. More recent studies demonstrated that a dimeric F_{AB} antibody fragments can also be exported by a hitchhiker mechanism via a signal peptide on only one of the two polypeptides that comprise the F_{AB} protein (DeLisa et al. 2003). In our approach, one protein is fused to a Tat signal peptide and the second protein to a protein reporter that can confer a phenotype only upon export into the periplasmic space. We demonstrated the detection of interacting proteins via the Tat two-hybrid system using two periplasmic reporters: maltose-binding protein, whose export can be detected either by selecting for growth on maltose or on MacConkey plates, and DsbA which catalyzes the formation of active AP in the periplasm. The two reporters allow the detection of protein interactions by growth on selective media, on indicator plates or by enzymatic assays. Both MBP and DsbA have been used as fusion partners to increase solubility, a feature that can be important for proteins that are susceptible to

aggregation. Furthermore, MBP fusions can be easily purified by one-step affinity chromatography on amylose resins thus expediting subsequent *in vitro* analyses of protein interactions.

Perhaps the most interesting and useful feature of the Tat two-hybrid assay is that it takes advantage of the requirement that proteins have to be folded in order to be competent for export via the Tat transporter. Neither the detailed mechanism nor the features of the folded polypeptide that are interrogated to determine export-competence are understood at the moment. However, Fischer et al. (2006) showed that Tat selects for proteins exhibiting higher solubility (Fisher et al. 2006). We therefore believe that the Tat two-hybrid assay is likely to be useful for reducing false positives that can arise from the non-specific interactions or aggregation of non-interacting polypeptides, as well as spurious positives arising from properties of a single partner (e.g. inadvertent activation/repression of transcription or stress responses that result in a signal).

During the early steps in their biogenesis, proteins destined for Tat export interact with signal sequence specific (e.g. TorD) or general chaperones such as DnaK (Graubner et al. 2007; Perez-Rodriguez et al. 2007). There is no evidence that these cytoplasmic chaperones are exported together with their target polypeptide. Consequently, weak interactions, such as those between chaperones and their substrates are not likely to be detected by the Tat 2-hybrid system. This may not be undesirable for studies aimed at finding specific binding partners The lower limit of protein affinities required to give a positive signal has not yet been determined.

The Tat 2-hybrid is expected to give false negatives, i.e. fail to detect interacting proteins, in two instances: First when the interacting protein partners form a complex

which is larger than the size limit that can be accommodated by the Tat pore (<180 kDa). Given that the DsbA reporter has a M.W. of 21 kDa this means that the total size of the complex formed by the interacting proteins must be less than 160 kDa. Second, false negative may arise when one of the two partner polypeptides exhibits folding kinetics that are substantially slower than the typical transit times for Tat export. When the folding kinetics are slow, the off pathway reactions leading to aggregation in the cytoplasm predominate and interfere with Tat export (Ribnicky et al. 2007).

In conclusion, the Tat two-hybrid assay could improve the accuracy of proteome wide two-hybrid analyses. In addition, a further useful biotechnological application for the Tat two-hybrid system will be in screening of libraries of mutant polypeptides to isolate clones that bind to a target protein with increased affinity or to increase solubility while still binding to its interaction partner.

Chapter 3

E. coli tatC mutations that suppress defective twin-arginine transporter signal peptides

3.1 INTRODUCTION

The bacterial Twin-Arginine Translocation pathway (Tat) is responsible for the transport of cofactor-containing proteins and other folded polypeptides across the cytoplasmic membrane. Proteins exported via Tat contain long signal peptides featuring a conserved twin-arginine motif with a consensus sequence of S/T-R-R-x-F-L-K (Berks 1996). In *Escherichia coli*, the functional Tat translocon is composed of the membrane proteins TatA, TatB and TatC. Among Gram-positive bacteria, a few organisms can be found containing only two identifiable *tat* genes, encoding TatA and TatC homologues and lacking a *tatB* gene (Ize et al. 2002; Jongbloed et al. 2002; Pop et al. 2002; Yen et al. 2002). In *E. coli* TatA is purified as a series of high molecular weight homo-oligomers that form pore-like structures of varying diameters. Based on these observations, Gohlke et al. (2005) proposed that TatA forms the protein conducting pore that assembles on demand, the size of which is determined by the dimensions of the translocated protein substrate (Gohlke et al. 2005). Biochemical analyses have revealed that TatB co-purifies in a complex with TatC and that this complex plays a key role in the targeting of Tat substrates to the translocon (Alami et al. 2003; Dabney-Smith et al. 2006; Gerard and Cline 2007). Based on cross-linking of *in vitro* translocated proteins into inverted membrane vesicles, Alami et al. (2003) discovered that TatC in particular exhibits

extensive contact with the signal peptide and recognizes the twin-arginine motif (Alami et al. 2003).

We sought to obtain genetic evidence that supports the role of TatC in the recognition of the twin-arginine motif and to further define the TatC epitopes that are important for this process. Substitution of the nearly invariant twin-arginine dipeptide by two lysines completely abolishes the export of all physiological Tat proteins tested so far (Buchanan et al. 2001; DeLisa et al. 2002) . However, there is evidence for very low level export of heterologous proteins such as colicin V(Ize et al. 2002) and the maltose binding protein (MBP) when fused to the TorA signal peptide in which the twin arginine motif had been mutated to a twin-lysine dipeptide (TorA(KK))(Kreutzenbeck et al. 2007b).

Here we report the isolation and characterization of TatC suppressor mutations that allow export of precursor proteins fused to a defective TorA(KK) signal peptide. These results provide additional genetic evidence that TatC serves as the receptor for Tat signal peptides. In addition, the suppressor mutations we have obtained help define regions in TatC that play a role in the recognition of the twin-arginine motif. As this manuscript was about to be submitted for publication, Freudl and coworkers (Kreutzenbeck et al. 2007b) reported the isolation of suppressor mutations in a plasmid encoding the TatABCE protein conferring export of a ssTorA-MBP fusions in which the RR dipeptide had been substituted by a lysine-glutamine sequence. Kreutzenbeck et al. (2007) presented evidence showing that the suppressor mutations permitted near wild type level of export of ssTorA-MBP containing a wild type RR dipeptide. In contrast, mutants we report here appear to represent a different class of suppressors that allow export of TorA(KK) signal peptides, but impair translocation of native Tat proteins.

3.2 MATERIAL AND METHODS

3.2.1 Bacterial strains and plasmids

Bacterial strains and plasmids are summarized in Table 3.1, their construction is described below.

Table 3.1. Bacterial strains and plasmids

Strain/plasmid	***Genotype***	***Reference***
MC4100	*F⁻ lacΔU169 araD139 rpsL150 thi flbB5301 deoC7 ptsF25 relA1*	(Casadaban and Cohen 1979)
B1lk0	MC4100, ΔtatC	(Bogsch et al. 1998)
GB22KK-C	MC4100, Δ*dmsABC*::Kan^R, *torA* R11K:R12K, Δ*tatC*::Spec^R	(Buchanan et al. 2002)
Xl-1 blue	F⁻ *thr leu thi lacY galK galT ara fhuA tsx dam dcm supE44*	Stratagene
pTGS	*sstorA-gfp-ssrA* in pBAD33	(DeLisa et al. 2002)
pKGS	*sstorA* (R11K, R12K)-*gfp-ssrA* in pBAD33	(DeLisa et al. 2002)
pBR322-*tatC*	*tatC* expressed from the tat promoter in pBR322	This study
pTatC-His8	pBR322-*tatC,* TatC with a C-terminal octahistidine tag	This study
pTatC-N6	pBR322-*tatC*, W92G, P97S P172L	This study
pTatC-W10	pBR322-*tatC*, P142S	This study
pTatC-W10-His8	pTatC-His8, P142S	This study
pTatC-X19	pBR322-*tatC*, F31S, P67L F75S, F94S,	This study
pTatC-Y2	pBR322-*tatC*, S2P, L99P, F124	This study
pTatC-W92G	pBR322-*tatC*, W92G	This study
pTatC-F94S	pBR322-*tatC,* F94S	This study
pTatC-F94S-His8	pTatC-His8, F94S	This study
pTatC-P97S	pBR322-*tatC,* P97S	This study
pTatC-W92G/P97S	pBR322-*tatC,* W92G, P97S	This study
pTatC-L99P	pBR322-*tatC*, L99P	This study
pTatC-L99P-His8	pTatC-His8, L99P	This study

3.2.2 Plasmids construction

tatC with its own promoter sequence was amplified from pFAT417 (Buchanan et al. 2002) using primers EMS21 (GCGGCGGACGTCTCTAGAGGAAGTGCAGCCGCAACTGG) and EMS22 (GCGGCGCCTGCAGGGTCGACTTATTCTTCAGTTTTTTCGCTTTCTGC) and then inserted into pBR322 between the AatII and SalI sites. A 0.6% nucleotide substitution

library of *tatC* was generated by error prone PCR (Fromant et al. 1995) using primers TatB2 (Buchanan et al. 2002) and EMS22; the mutagenized gene product was used to replace *tatC* gene in the pBR322-*tatC* construct. For immunostaining, an octahistidine tag was added by PCR amplification with the primers TatB2 and EMS23 (CGCCGCGTCGACTTAATGGTGATGGTGATGGTGTTCTTCAGTTTTTTCGCTTT CTGC). Point mutations were generated by site-directed mutagenesis using overlap PCR.

3.2.3 Growth conditions

All cultures were grown at 37˚C. For genetic manipulations, cells were typically grown aerobically in LB media overnight. For SDS growth plate analysis, 1.5% LB agar plates were supplemented with 2% SDS and appropriate antibiotics. For anaerobic TorA growth selection, cells were grown in closed tubes using Cohen and Rickenberg medium (Cohen and Rickenberg 1956) with 0.4% TMAO; glucose was replaced with 0.6% glycerol. For TMAO activity gels, cells were grown for 22-24 h in closed 15 ml Falkon tubes using Cohen and Rickenberg medium (CR medium) as previously described (Sawers et al. 1985) with 0.4% glucose and 0.4% TMAO (Sigma-Aldrich).

3.2.4 Flow cytometry

Cells were cultivated at 37˚C in LB after subculturing 1/20 from saturated overnight cultures. Cultures were induced in mid-logarithmic growth phase with 0.02% arabinose, and after a 5 h induction period, cells were washed once with PBS, diluted 100 times and analyzed by flow cytometry on a BD FACSort™ Cell Sorter or FACSAria. Successful sorting gates were set slightly above background (2 fold) based on fluorescence at 488 nm/530 nm.

3.2.5 Subcellular localization

15 ml cell cultures were normalized to a total O.D of 5, and then fractionated after the osmotic shock protocol (Stanley et al. 2000). Cells were normalized based on OD_{600} to a total of OD_{600} 5 and then fractionated by the cold osmotic shock procedure (Stanley et al. 2000). Briefly, cells were resuspended in fractionation buffer (30 mM Tris-HCl pH 8.0, 20% (w/v) sucrose, 1 mM EDTA pH 8.0) incubated at room temperature for 10 min, and then after centrifugation re-suspended in ice-cold $MgSO_4$ (5 mM). Spheroplasts were sonicated. Debris was removed by centrifugation.

3.2.6 TMAO activity gels

For TMAO activity gels, cells grown in CR medium with 0.4% glucose were fractionated and 12.5% of the obtained fractions were separated by polyacrylamide gel electrophoresis (PAGE) under native condition on a 7.5% Tris gel at pH 6.5. Gels were incubated with 20 ml of 100 mM sodium phosphate buffer (pH 7), 1 mM methyl viologen was added and reduced by an excess of a 1% dithionite solution in 0.1 mM NaOH. Gels were incubated under mildly shaking until stained blue; 10 ml of a 2 M TMAO in the same phosphate buffer was added to visualize TorA activity as it can be seen in the transparent lanes.

3.2.7 Western blot analysis

Western blotting was performed as described earlier (Levy et al. 2001). After PAGE separation, proteins were transferred to a methanol-activated polyvinylidine fluoride (PVDF) membrane (Millipore or Pall Life Sciences) by a semi-dry blotting

device (Owl), as recommended by the manufacturer. Membranes were blocked in 5% nonfat dry milk/TBST buffer (15 mM Tris-HCl, 4 mM Tris, 137 mM NaCl, 0.1% Tween-20, to pH 7.6) for 1 hour at room temperature. Anti-His antibodies (Sigma) were applied in a 1:5000 dilution in 1% milk PBST buffer overnight at 4°C. Anti-DsbA (1:10 000) and anti-GroEL (1:25 000) were incubated in 1% milk TBST for 1 h at room temperature. Anti-His antibody was detected with anti-mouse HRP conjugated antibody (BioRad) in a 1:10 000 dilution whereas the anti-DsbA antibody (Beckwith lab) and the anti-GroEL (Sigma) were detected with an anti-rabbit HRP conjugated antibody in a 1:10 000 dilution with 1% milk TBST. Presence of HRP was detected with the Supersignal® Western Pico substrate kit (PIERCE).

3.3 RESULTS

3.3.1 Isolation of suppressor TatC variants by FACS

ssTorA fused to the fluorescent reporter GFP-SsrA allows the facile and quantitative detection of Tat export by flow cytometry (DeLisa et al. 2002). The ssTorA-GFP-SsrA fusion is encoded by the plasmid pTGS (DeLisa et al. 2002) under the control of the arabinose promoter. The presence of the SsrA tag on GFP ensures that any fusion protein that fails to be exported via the Tat pathway is degraded by ClpXP (Levchenko et al. 2000). Thus, cell fluorescence arises only from ssTorA-GFP-SsrA that has been sequestered away from the proteolytic machinery following Tat-dependent export into the periplasmic space (DeLisa et al. 2002; Tullman-Ercek et al. 2007), see Figure 3.1).

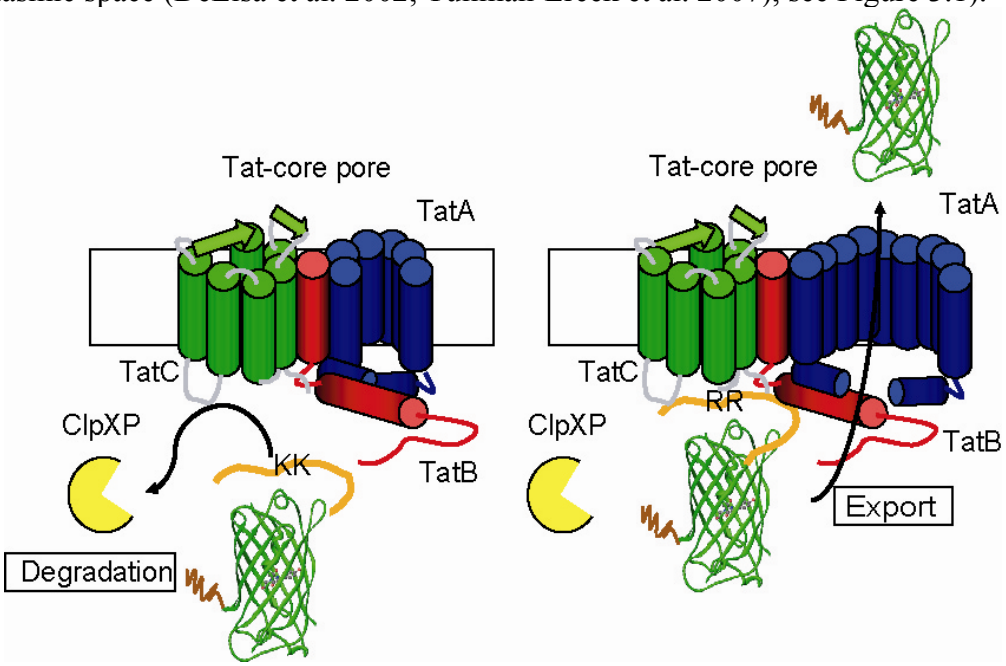


Figure 3.1. Schematic analysis of the screen for the isolation of suppressors.

E. coli B1lk0 (Δ*tatC*) (Bogsch et al. 1998) cells carry a complete deletion of the *tatC* gene and therefore exhibit background fluorescence with GFP-SsrA; in addition they are sensitive to anionic detergents such as SDS and form filaments. B1lk0 cells were transformed with the plasmid pBR322-*tatC* which encodes the *tatC* gene with its own ribosome binding site downstream of the constitutive *tat* promoter. B1lk0 pBR322-*tatC* cells also containing pTGS exhibited a 7-fold increase in GFP-SsrA fluorescence, normal cell morphology and were resistant to SDS (Figure 3.3 and data not shown). Mutation of the RR dipeptide within ssTorA-GFP-SsrA to a twin-lysine (KK) sequence reduced the cell fluorescence to the background level, but as expected did not have any effect on cell morphology or SDS resistance (Figure 3.3 and (DeLisa et al. 2002)).

The *tatC* gene was subjected to random mutagenesis by error prone PCR (Fromant et al. 1995). The resulting DNA was ligated into pBR322-*tatC* in place of the w.t. *tatC* gene. The ligation product was transformed into XL-1 blue cells (Stratagene) resulting in a library of $2x10^6$ independent transformants. The plasmid preparation of mutagenized pBR322-*tatC* was then introduced into *E. coli* B1lk0 cells transformed with a plasmid encoding ssTorA(KK)-GFP-SsrA (pKGS). Sequencing of 6 clones selected at random revealed a mutation rate of 0.6% nucleotide substitutions per gene.

B1lk0 cells expressing the *tatC* alleles and ssTorA(KK)-GFP-SsrA were then screened by flow cytometry and clones with increased fluorescence were isolated. No clones with fluorescence comparable to the wild type (around 7-fold higher than background) were obtained even after repeated attempts. However, four clones (B1lk0 with pTatC-N6, pTatC-W10, pTatC-X19 and pTatC-Y2) that reproducibly exhibited about 2-2.5 fold higher fluorescence above background were obtained and studied further (see Figure 3.2 for sequence alignment and Figure 3.3 for fluorescence).

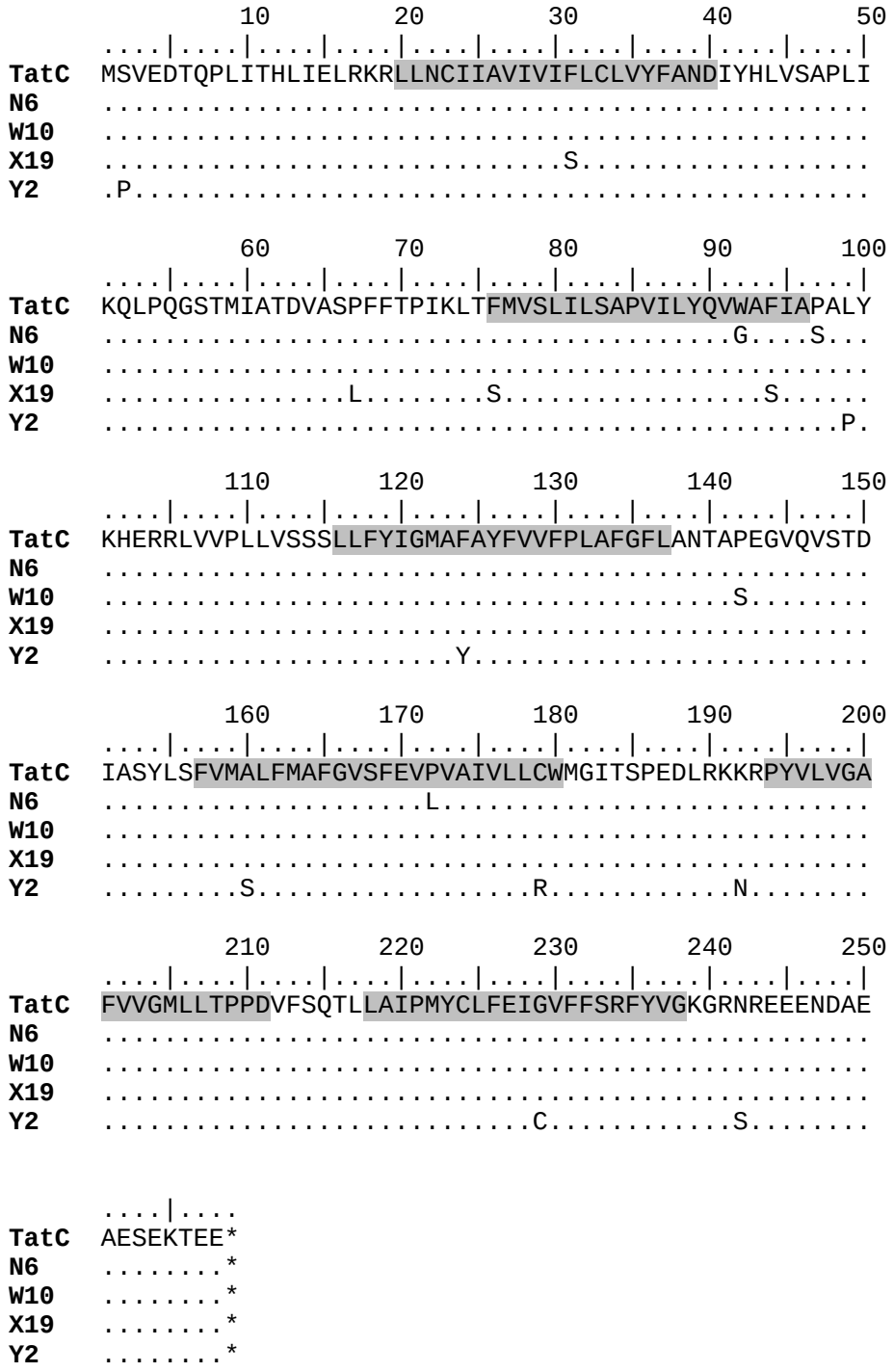
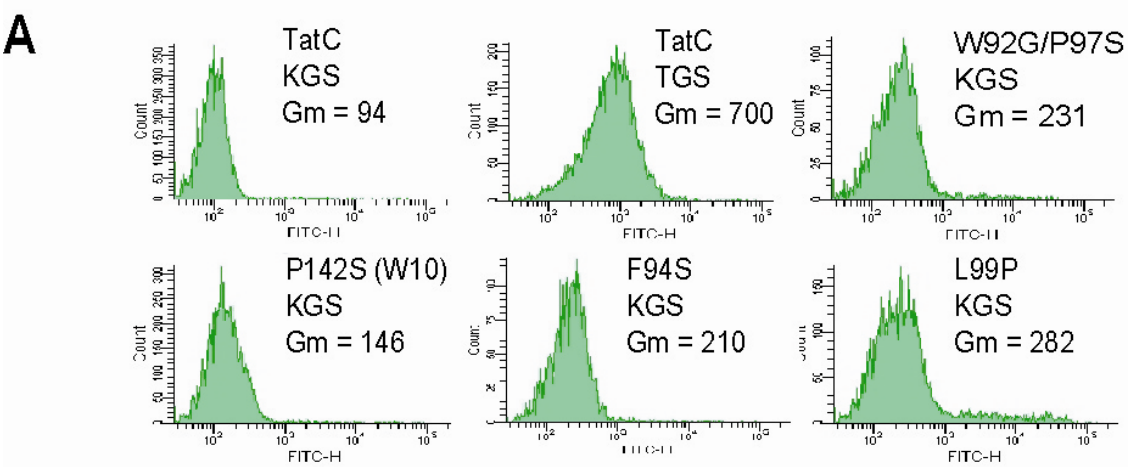


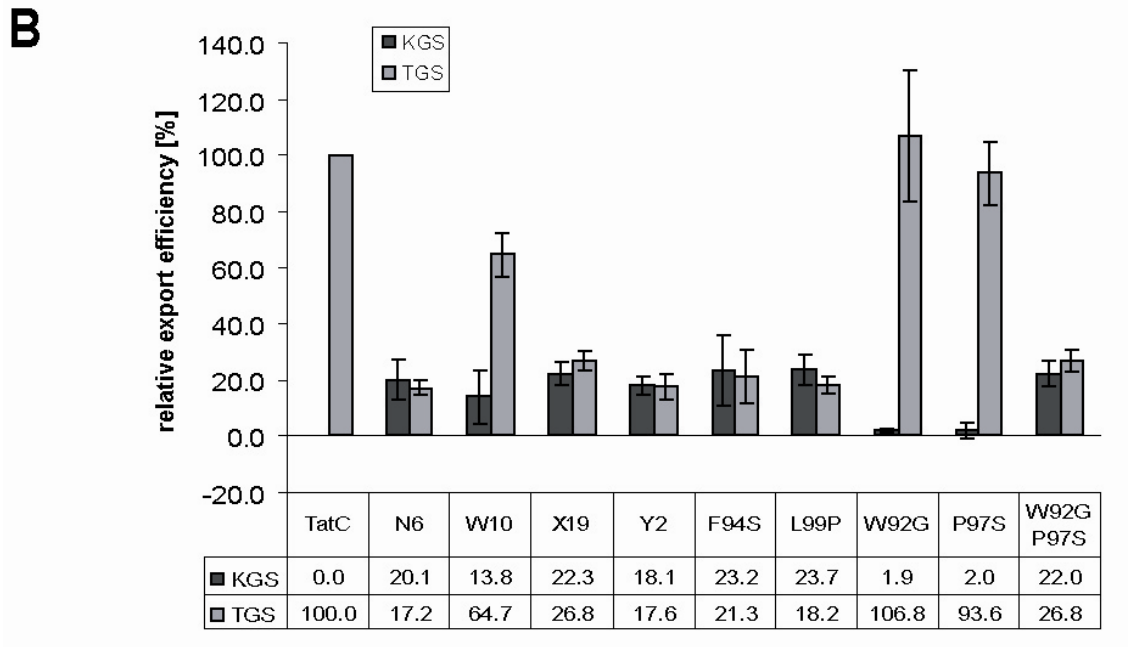
Figure 3.2. Sequence alignment of isolated variants. Grey regions indicate predicted transmembrane domains after Buchanan et al.(Buchanan et al. 2002) and Ki et al.(Ki et al. 2004).

The elevated fluorescence phenotype was dependent on the expression of both the TatC variant and the ssTorA(KK)-GFP-SsrA reporter protein, suggesting that the mutations act as suppressors of the export defect caused by the KK substitution in the ssTorA signal peptide.

With the exception of pTatC-W10, which contained the single amino acids substitution P142S, all isolated *tatC* genes contained multiple amino acid changes. Several of the respective single amino acid mutations (W92G, F94S, P97S, L99P, A160S, V167D, L172P, C179R and K192N) as well as the double amino acid substitution W92G/P97S were generated by site specific mutagenesis and their ability to export ssTorA(KK)-GFP-SsrA was analyzed by FACS. TatC carrying the amino acid substitutions F94S and L99P conferred cell fluorescence comparable to clones pTatC-X19 and pTatC-Y2, respectively. The B1lk0 carrying pTatC-N6 contained the amino acid substitutions W92G, P97S and P172L. While neither the amino acid change in pTatC-W92G nor pTatC-P97S displayed export of ssTorA(KK)-GFP-SsrA above background, the double substitution variant W92G/P97S exhibited a fluorescence profile comparable to that of pTatC-N6 (Fig.3.3.B). None of the other tested single amino acid substitutions conferred fluorescence above background in cells co-transformed with pKGS. To investigate the effect of mutations on the substrate specificity for the w.t. TorA signal peptide, we analyzed the fluorescence of the mutants expressing ssTorA-GFP-SsrA (Fig. 1.C.). We calculated the relative export efficiency after subtracting the background fluorescence of B1lk0 cells transformed with pBR322-*tatC* and pKGS. The export efficiency of the w.t signal peptide fusion for all suppressing variants except for the clone carrying pTatC-W10 (containing the single mutation P142S) was around the same amount as for ssTorA(KK)-GFP-SsrA, indicating a partial impairment of export. The

TatC P142S variant conferred around 65% transport efficiency of ssTorA-GFP-SsrA compared to that of the w.t. TatC plasmid.





	TatC	N6	W10	X19	Y2	F94S	L99P	W92G	P97S	W92G P97S
KGS	0.0	20.1	13.8	22.3	18.1	23.2	23.7	1.9	2.0	22.0
TGS	100.0	17.2	64.7	26.8	17.6	21.3	18.2	106.8	93.6	26.8

Figure 3.3. Fluorescence histograms of *E. coli* Blk0 cells expressing either ssTorA-GFP-SsrA (pTGS) or ssTorA(KK)-GFP-SsrA (pKGS). (A) Cultures were induced in mid-logarithmic phase with 0.02% arabinose, and after a 5 h induction period, cells were washed once with PBS, diluted 100 times and analyzed by flow cytometry on a BD FACSort™ Cell Sorter or FACSAria. The fluorescence histograms of single amino acid suppressors and a reconstructed double amino acid substitution exporting ssTorA(KK)-GFP-SsrA (KGS) are indicated, data shown was acquired by FACSAria. (Gm: geometric mean fluorescence, TGS: ssTorA-GFP-SsrA; KGS: ssTorA(KK)-GFP-SsrA). (B) Average relative export efficiencies of isolated mutants and single point mutations exporting ssTorA(KK)-GFP-SsrA and the wildtype TorA signal peptide fused to GFP-SsrA. The geometric mean fluorescence of the negative control (B1lk0 pBR322-*tatC* with pKGS) was subtracted and the results normalized by comparing the values obtained to the fluorescence in B1lk0 pBR322-*tatC* expressing pTGS. Relative export efficiency was the average of three independent measurements from 3 separate cultures for each clone; error bars represent the standard deviation of the relative export efficiency.

The level of suppression conferred by the various single amino acid TatC variants may also be related to the expression level of the respective proteins. In pBR322-*tatC*, *tatC* expression occurs from the weak *tat* promoter and a pBR322 origin. Attempts to detect TatC produced from pBR322-*tatC* using Western blot analysis with existing polyclonal sera (Buchanan et al. 2002; de Leeuw et al. 2002; Punginelli et al. 2007) were unsuccessful, evidently because of the low level of *tatC* expression. To circumvent this problem TatC was fused to an octahistidine tag (His8). Consistent with previous observations we confirmed that a C-terminal His8 fusion does not interfere with the function of w.t. TatC or with that of the suppressor variants (data not shown and (Buchanan et al. 2002; de Leeuw et al. 2002)). Western blot analysis of TatC variants in total cell lysate revealed that the expression level of pTatC-F94S-His8 and pTatC-W10-His8 (P142S) was identical to that of w.t. pTatC-His8 expressed whereas pTatC-L99P-His8 accumulated at somewhat lower levels (Figure 3.4). Thus at least for 2/3 of the

single amino acid TatC variants, the degree of suppression does not stem from differences in expression level.

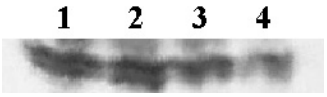


Figure 3.4 Western blot analysis of TatC and single amino acid variants. For immunodetection of the expression of the *tatC* single point mutation suppressor variants, containing a C-terminal ocathistidine tag. B1lk0 cells transformed with the above plasmids were grown in LB media overnight, harvested and the same number of cells normalized by $O.D._{600}$, were loaded (1) pTatC-P142S-His8, (2) pTatC-His8, (3) pTatC-F94S-His8, (4) pTatC-L99P-His8 expressed in B1lk0.

3.3.2 Monitoring activity of native TorA enzyme

Further, we examined whether the putative suppressors identified in the screen above could export a native *E. coli* protein having a twin-lysine substitution in the RR dipeptide. The enzyme trimethylamine N-oxide reductase (TorA) reduces trimethylamine N-oxide (TMAO) to trimethylamine. TorA contains a bismolybdopterin guanine dinucleotide cofactor, which is incorporated into the apo-protein within the cytoplasm, a process that requires the chaperone TorD (Pommier et al. 1998; Ilbert et al. 2003). Following cofactor incorporation, the TorA holoenzyme is exported into the periplasm by the Tat machinery (Santini et al. 1998). Reduction of TMAO has to occur on the periplasmic side of the inner membrane in order to assure electron transfer from TorC, to TorA onto TMAO (Gon et al. 2001).

In *E. coli* GB22KK-C (*torA* R11K,R12K; Δ*dmsABC*::Kan^R, Δ*tatC*::$Spec^R$) the twin-arginine dipeptide in the signal peptide of the chromosomal *torA* gene has been replaced with a KK sequence (Buchanan et al. 2001) and there is a complete *tatC* deletion marked with a spectinomycin resistance cassette (Buchanan et al. 2002). Strain

GBKK22-C also lacks *dsmABC* gene that encodes the only other enzyme that is known to reduce TMAO (Buchanan et al. 2001). The TorA protein in GBKK22-C cannot be exported, but it is competent for cofactor assembly and as a result, the active enzyme accumulates in the cytoplasm (Buchanan et al. 2001). GBKK22-C cells expressing the TatC variants encoded by pTatC-N6, pTatC-W10 (P142S), pTatC-X19, pTatC-Y2, the TatC single amino acid substitution construct pTatC-L99P or the double amino acid substitution pTatC-W92G/P97S, were grown anaerobically for 22-24 h in Cohen and Rickenberg medium (Cohen and Rickenberg 1956), harvested, and fractionated by the osmotic shock procedure(Stanley et al. 2000). The efficiency of the fractionations was evaluated by monitoring the level of GroEL and DsbA as cytoplasmic and periplasmic markers, respectively. No GroEL could be detected by Western blot analysis in the periplasmic fractions whereas DsbA was localized in the periplasmic fraction as expected (data not shown). Periplasmic and cytoplasmic proteins were then resolved by native electrophoresis and the gel was stained with methyl viologen in the presence of the reducing agent dithionite. Reduced methyl viologen served as an artificial electron donor for TMAO reductase resulting in a readily detectable color change from blue to transparent upon oxidation of TMAO by the mature TorA enzyme (Silvestro et al. 1988) (Figure 3.5). In *E. coli* MC4100 cells, about 50% of the active TMAO reductase was localized in the periplasmic fraction. No periplasmic TMAO reductase band was detected in *E. coli* GBKK22-C expressing wildtype TatC from pBR322-*tatC* (Figure 3.5). Enzymatically active TMAO reductase, was clearly detected in periplasmic fractions of cells co-expressing the TatC suppressors plasmids except for cell with pTatC-W10 (Figure 3.5).

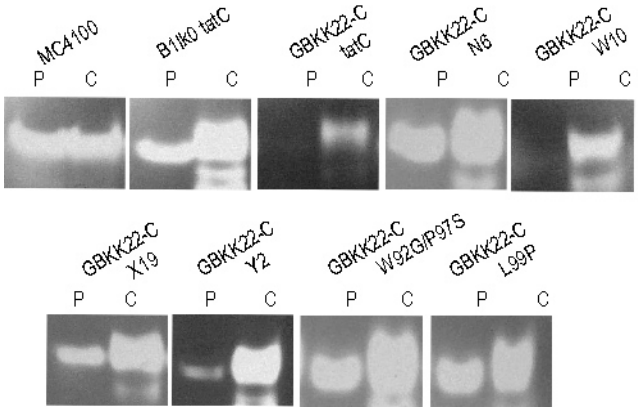


Figure 3.5. TMAO activity gels. Gels were stained with reduced methyl viologen. Reduced methyl viologen serves as artificial electron donor for the reduction of TMAO by TMAO reductase, resulting in transparent lanes. MC4100 and B1lk0 (Δ*tatC*) containing pBR322-*tatC* were used as positive controls. GB22KK-C cells contain the a gene encoding TorA(R11K,R12K) in the chromosome and the deletion of *dmsABC*::*kan* (Buchanan et al. 2002). All other variants are indicated.

However, we noticed that cells expressing the plasmid pTatC-W10 carrying the *tatC* P142S allele grew anaerobically with glycerol (0.6%) and TMAO (0.4%) as the final electron acceptor. Since periplasmic TMAO reductase activity is required for growth on TMAO, TatC P142S variants must be capable of exporting TorA having a KK mutation in its signal sequence.

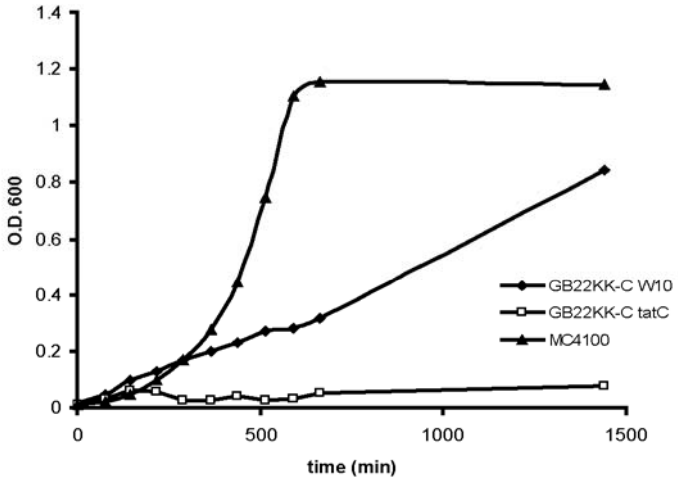


Figure 3.6. Anaerobic growth of GB22KK-C cells expressing the *tatC* W10 (P142S) variant. Non-fermentable glycerol was used as the carbon source, forcing the bacterium to utilize TMAO as final electron acceptor. Overnight cultures were diluted 1/100 in Cohen and Rickenberg medium (Cohen and Rickenberg 1956) with 0.6 % glycerol and 0.4% TMAO (Sigma), and grown in 1.5 ml screw cap vials filled to the top and incubated at 37˚C.

Consequently, it is possible that the failure to detect soluble TorA in the periplasmic fraction of these cells may be due to the retention of the enzyme on the surface of the spheroplasts. Alternatively it is possible that a very low level of exported TMAO that cannot be detected by the assay we used may still be sufficient to allow growth under these conditions.

3.3.3 Effects of *tatC* suppressor alleles on SDS resistance and cell separation

In *E. coli,* the N-acetylmuramoyl-l-alanine amidases AmiA and AmiC are exported by Tat (Ize et al. 2002). These enzymes cleave the peptide moiety of the N-acetylmuramic acid in peptidoglycan. Mutations that block the Tat pathway and prevent the export of AmiA and AmiC result in sensitivity to SDS (Ize et al. 2002) and partially inhibit cell division (Heidrich et al. 2001). In *amiC*-deficient strains, the daughter cells separate poorly and as a result up to 30% of the population forms chains (Bernhardt and de Boer 2003). *amiA* mutants exhibit a milder defect in septation with only 5-10% of the cell population forming short chains. *amiA* and *amiC* double mutants show a much more pronounced cell chain phenotype than *amiC* mutants (Heidrich et al. 2001; Bernhardt and de Boer 2003; Priyadarshini et al. 2006). Light microscopy of stationary phase cells expressing the *tatC* suppressor genes in B1lk0 (Δ*tatC*) revealed different degrees of cell chain formation (Fig 3.7). The TatC P142S variant gave the least defect in cell separation with several short chains (3-4 cells) observed in addition to a few single cells. B1lk0 cells containing pTatC-N6 showed a higher percentage of longer chains (>4 cells/chain). The TatC W92G/P97S substitutions that recapitulate the cell fluorescence and TMAO activity phenotypes of cells carrying pTatC-N6 which contains these two amino acid changes next to the substitution P172L, showed a lower percentage of long chains. The TatC variants, encoded by pTatC-X19, pTatC-Y2, pTatC-F94S and pTatC-L99P showed severe impairment in cell division. Collectively, this data indicates that none of the suppressor mutations completely restored a normal cell morphology suggesting that normal Tat export of proteins involved in cell division, presumably AmiA and AmiC, is at least partially impaired.

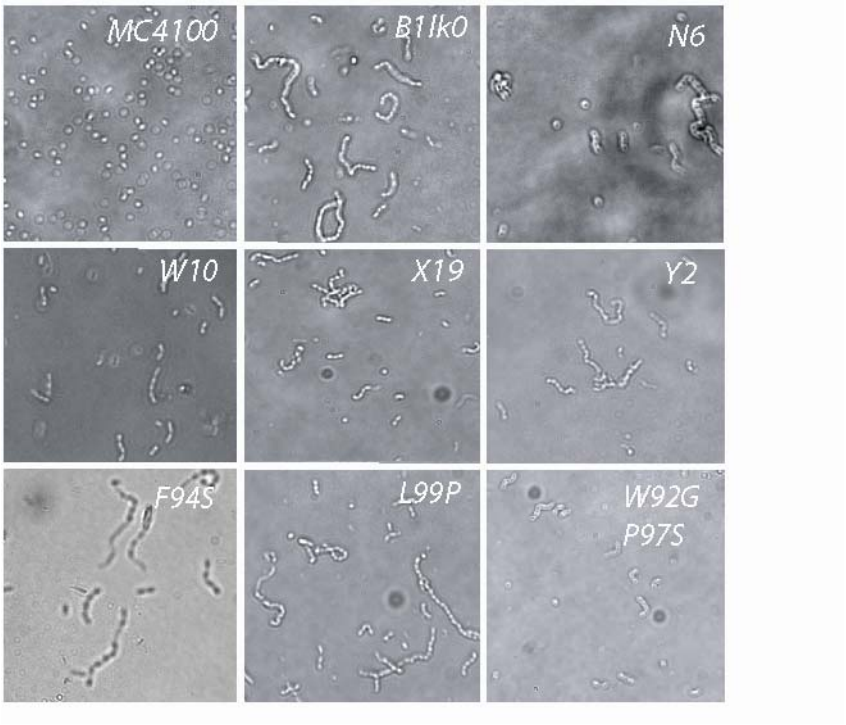


Figure 3.7. Morphology of TatC suppressors. Light microscopy photographs of stationary phase of B1lk0 transformed with plasmids expressing the indicated *tatC* suppressor alleles. Cells were grown in LB medium at 37˚C.

A functional Tat export system is required for normal cell envelope formation as determined by the ability of the cells to grow in the presence of the anionic detergent SDS. W.t. *E. coli* cells are almost unaffected by the presence of 2% SDS. The *tatC* strain B1lk0 fails to form colonies on agar plates with 2% SDS, but upon transformation with pBR322-*tatC* completely restored SDS resistance. However, cells expressing the *tatC* suppressor alleles failed to form colonies on LB plates with 2% SDS. The only exception was the TatC P142S substitution which showed normal colony formation on 2% SDS (data not shown). These observations provide further evidence that the TatC

suppressor mutations do not support the export of the native Tat proteins which are responsible for detergent resistance.

3.4 DISCUSSION

Substitution of the prototypical twin-arginine sequence found in Tat signal peptides by a lysine dipeptide prevents translocation via the Tat pathway in all organisms examined so far, presumably because this mutation inhibits the interaction of the signal peptide with *tatC* (Alami et al. 2003). A search for mutations in the *tatC* gene that allow export of ssTorA(KK)-GFP-SsrA, conferring increased cell fluorescence, ultimately led to the isolation of two different types of suppressor *tatC* alleles having a single amino acid substitution each. Two single amino acid suppressor mutations located on the cytoplasmic site of TatC conferred 20-25% of the cell fluorescence relative to the level observed with the w.t. *tatC* regardless of the signal peptide (RR or KK) used. We also identified two mutations, W92G and P97S that acted in an epistatic manner to suppress the export defect resulting from the substitution of the RR dipetide by KK. Neither of these two amino acid substitutions alone conferred export of signal peptides with a KK dipeptide. In addition, the two single amino acid suppressors and the W92G/P97S allele were shown to confer export of TMAO reductase (TorA) fused to a signal peptide containing the KK substitution. Finally, the cells displayed defects in septation and in SDS sensitivity. On the other hand, the amino acid substitution P142S, constitutes a different kind of suppressor in that it confers around 15% relative export efficiency for ssTorA(KK)-GFP-SsrA, but a significantly higher relative export efficiency of the ssTorA-GFP-SsrA (~65%). This finding and also the cell's higher resistance to SDS suggest that export of Tat substrates containing w.t. signal peptides is not greatly

impaired by the TatC P142S mutation. While we could not detect TMAO activity in the periplasmic space of GB22KK-C with the TatC P142S variant encoded by pTatC-W10 the finding that these cells can grow anaerobically with TMAO as a final electron acceptor, a process that depends on the presence of TMAO reductase in the periplasm, revealed that export of the enzyme was taking place.

The topology of TatC within the membrane has been determined experimentally using numerous reporter fusions (Drew et al. 2002; Behrendt et al. 2004; Ki et al. 2004). According to this model, the first type of suppressor mutations, F94S, L99P and W92G/P97S, are located at the end of the second transmembrane domain and at the beginning of the first cytoplasmic loop of TatC, respectively. The first cytoplasmic loop, also referred to as the 2nd cytoplasmic domain, of TatC corresponds to a highly conserved region in the bacterial and chloroplast TatC family and it plays a critical role in Tat export (Allen et al. 2002; Buchanan et al. 2002). In earlier studies it was reported that the mutation L99A severely reduces the export of authentic Tat substrates to less than 5%. However, the effect of this mutation with respect to KK export had not been tested (Buchanan et al. 2002). Similarly it has been reported that an F94A substitution completely abolishes export.

The mutation P142S is predicted to be located in the second periplasmic loop raising the question of how it is able to affect the recognition of the signal peptide since this event is expected to take place, or at least be initiated in, the cytoplasm. However, in a search for *tatC* mutants that allow export of a defective ssTorA(KQ)-MBP fusion, Kreutzenbeck et al. (2007) reported the isolation of the amino acid substitution E8K located at the N-terminus of TatB, a domain which is predicted to be localized in the periplasm. The finding that mutations in periplasmic residues in TatC or in TatB can impact the recognition of the signal sequence highlights the complexity of the

organization of the Tat translocon. As it has been noted by Kreutzenbeck et al. (2007), mutations in residues exposed to the periplasmic side could result in a re-organization of the tertiary structure of TatC or TatB. For the thylakoid Tat pathway, signal peptide binding was reported to occur in a binding pocket deep inside of the membrane domains of the TatB/TatC (Hfc106/cpTatC) receptor complex (Gerard and Cline 2007). Thus, it is conceivable that mutations in a periplasmic domain may affect recognition events that occur either deep within the membrane or perhaps in proximity of the periplasmic site. Alternatively, we cannot rule out the possibility that during the translocation cycle TatB and also some domains of TatC may undergo dramatic conformational changes, as it has been proposed to be the case for TatA (Gouffi et al. 2004).

The *tatC* mutations we identified in the present study exhibit different phenotypes relative to those reported recently by Kreutzenbeck et al. (2007). The latter were isolated based on their ability to confer growth of cells expressing ssTorA(KQ)-MBP on maltose plates but they were also shown to allow efficient recognition or even higher membrane translocation rates of w.t. ssTorA-MBP. We utilized a more stringent screen in which export could be quantitatively determined based on cell fluorescence, and isolated mutants that suppressed the export defect of ssTorA(KK)-GFP-SsrA. We found two different types of suppressors: One type of suppressor variants exhibited a ~80% reduction in the export of ssTorA-GFP-SsrA and also reduced ability to export authentic Tat substrates as manifested by septation defects and SDS sensitivity. These suppressors mutations, namely substitutions F94S, L99P and W92G/P97S, are located in the first cytoplasmic loop of TatC, unlike the majority of Kreutzenbeck's mutations, which localized within the first few N-terminal residues of the protein. In addition, we isolated the mutation P142S which is located within a putative periplasmic loop and has a more

moderate effect on the export of authentic substrates, thus leading to SDS resistance but partial defects in septation.

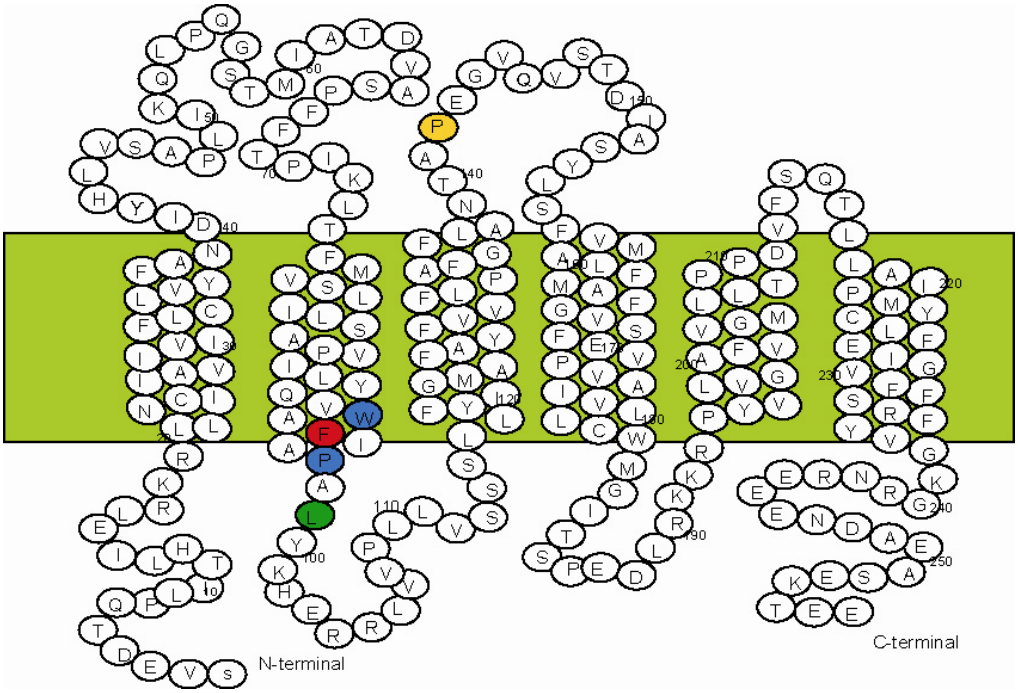


Figure 3.8. Predicted location of the TatC suppressor mutations. The topological model of TatC was obtained from Buchanan et al.(Buchanan et al. 2002), Behrendt et al.(Behrendt et al. 2004) and Ki et al.(Ki et al. 2004)

The suppressors we characterized in the present study, F94S, L99P, W92G/P97S and P142S, lie in the center of the protein sequence at opposite sides of the third transmembrane domain (Figure 3.8). All suppressor mutations resulted in non-conservative substitutions that are likely to have a significant effect on protein conformation. Interestingly, the TatC single point mutations identified by Kreutzenbeck et al. (2007) and the ones reported in the present study were all located within the N-

terminal half of the protein. This suggests that the N-terminal of TatC is likely to be particularly important for the recognition and interactions with the signal peptide. In agreement with these observations recent *in vitro* cross-linking studies also highlighted the significance of the N-terminus of TatC for signal peptide recognition (Holzapfel et al. 2007).

As was the case with the mechanistic elucidation of the Sec export pathway, the synergism between genetic and biochemical analysis is expected to lead to deep insights that would be unlikely to arise solely by either approach.

Chapter 4

Studies on the display of GFP variants and randomized GFP selections via a Tat-based phage display system

4.1 INTRODUCTION

Antibodies bind tightly and with high specificity to foreign molecules. Their species-specific core part recruits other modules of the immune response to eliminate the intruding molecule or microorganism from the host. It is because of their high specificity that antibodies are used for a variety of applications in research. Proteins that bind with high specificity are not only useful for diagnostics, but are also of interest for drug development and protein purification.

The green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* has been widely utilized as a reporter for gene expression (Chalfie et al. 1994), protein sorting and localization (Kaether and Gerdes 1995; Santini et al. 1998; DeLisa et al. 2002); and it has been further used in analyses of organelle shape and mobility (Rizzuto et al. 1995; Cole et al. 1996) as well as membrane protein topology (Drew et al. 2002; Drew et al. 2005) and to monitor diffusion of proteins within membranes or the lumens of organelles (Cole et al. 1996; Subramanian and Meyer 1997). GFP has a β-barrel structure. The loops joining the b-sheets resemble the complementary-determining regions (CDR) of the variable loops of antibodies (Figure 4.1.A and B). Hence it is conceivable that the surface exposed loops of GFP could provide an analogue to the antigen recognition regions of antibodies. The advantage of having a GFP variant with intrinsic binding affinity for a protein of interest (POI) is that no invasive technique or detection with substrates or co-

factors will be required to monitor its presence. An antigen-binding GFP variant therefore would be a powerful biosensor that can be further endogenously expressed in a broad range of host organisms (Cubitt et al. 1995).

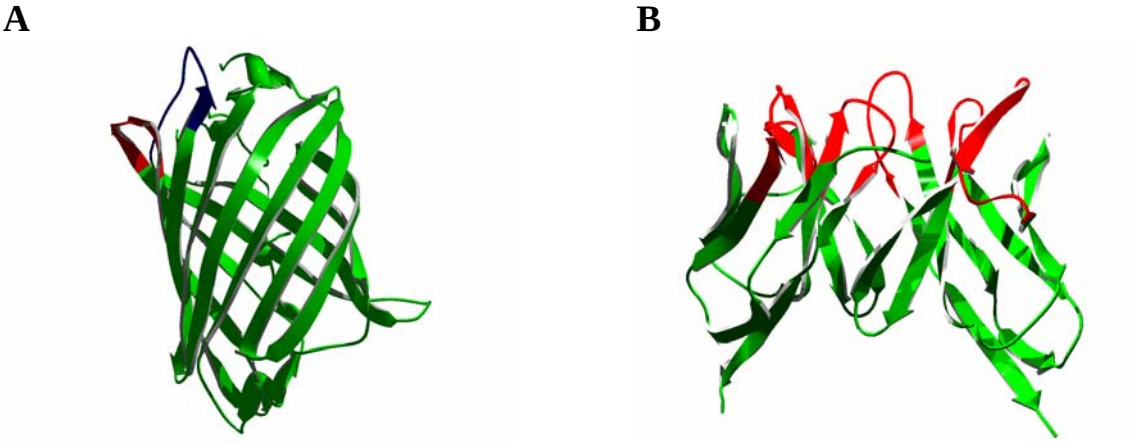


Figure 4.1. (A) Loop containing Q157 and K158 marked in red. In blue, residues 190-200 which can be potentially replaced by a random peptide stretch. (B) M18 scFv with marked binding loops.

A variety of display systems are available for the isolation of proteins that bind ligands from large combinatorial libraries of protein (Daugherty 2007). By far the most commonly used display one is the filamentous phage display technology (Smith 1985). However, phage display is dependent on protein secretion through the Sec pathway which is incompatible with GFP. When exported via Sec GFP is translocated into the periplasm where it folds into a soluble form, but the chromophore is not established (Feilmeier et al. 2000). We therefore decided to utilize a Tat-based phage display technology which allows the export of the POI by the Tat pathway (Figure 4.2). Since pIII is incompatible with Tat export, it is targeted to the Sec pathway. Association of the POI with the pIII anchor protein is mediated by the two adaptor proteins Jun and Fos, one of each fused to the POI and pIII (Paschke and Hohne 2005), Tullman-Ercek,

unpublished results). A Tat-based phage display has the major advantage that it allows the POI to fold in the cytoplasmic environment of the bacterial cell where cytoplasmic folding chaperones such as GroEL, ClpB, DnaK/DnaJ/GrpE are available. Translocation via Tat further implies that folding proof-reading mechanism is applied to the POI which ensures that only correctly folded proteins are exported (DeLisa et al. 2003) and displayed.

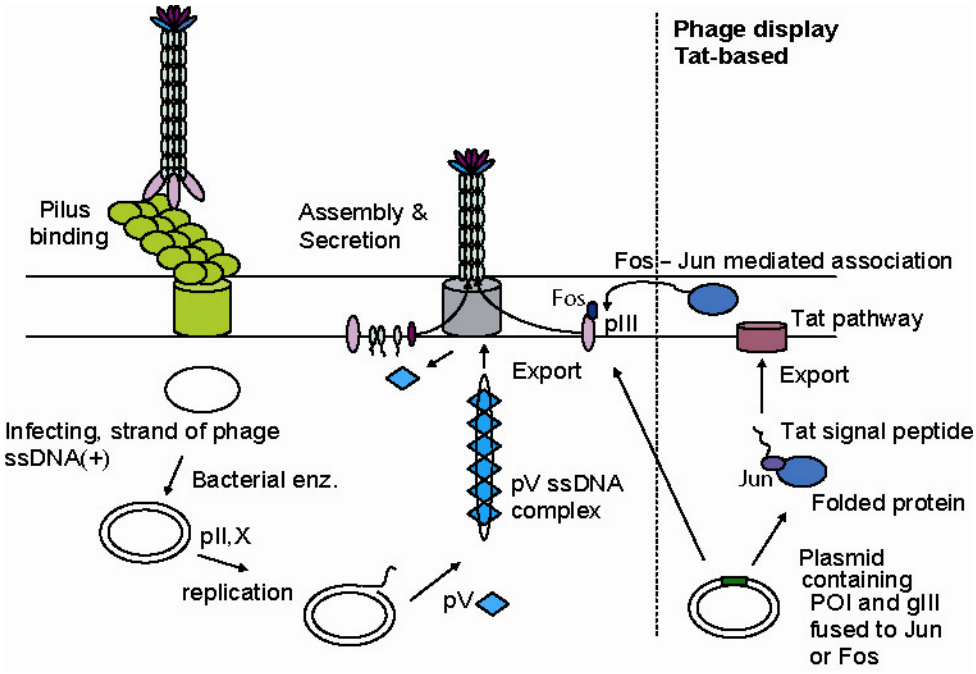


Figure 4.2. Tat-based phage display. The protein of interest (POI) fused to Jun and after folding in the cytoplasm, it is exported via the Tat pathway. The pIII protein is fused to Fos and is exported via the Sec pathway. Association of the POI with the anchoring pIII protein is mediated by the interaction of Jun and Fos in the periplasm. The interaction of Jun and Fos is stabilized by disulfide bridges. The pIII chimera is then anchored into the newly assembled phage. The life cycle of the phage is not affected except for the incorporation of a chimeric pIII protein. The phage particle associates with the F-pilus of the *E. coli* cell. The single stranded phage genome is inserted into the cytoplasm where it recruits bacterial enzymes to assist its replication. The phage proteins are translated and its own secretion machinery is assembled which allows the secretion of the newly generated phage particle. In this system the phage particle has w.t. pIII and chimeric pIII protein, as it is typical for a 3+3 phage vector system.

4.2 MATERIAL AND METHODS

4.2.1 Bacterial strains and plasmid

For plasmid propagation, cells were typically grown in LB media at 37˚C. For phage amplification, cells were grown in 2xYT as described in the methods sections below. Plasmids are summarized in Table 4.1 and Figure 4.2.

Table 4.1. Bacterial strains and plasmids

Name	***Genotype***	***Source***
DH5α	F-(f80dlacZΔM15,) Δ(lacIZYA-argF)U169 deoR recA1 endA1 hsdR17(rk-,mk+) supE44, thi-1 gyrA96, relA1	Lab strain
DHF	(f80dlacZΔM15,) Δ(lacIZYA-argF)U169 deoR recA1 endA1 hsdR17(rk-,mk+) supE44, thi-1 gyrA96, relA1, (F' proAB, lacIq, lac z ΔM15, Tn10 (tetR)	This study
MC1061	hsdR2 hsdM+ hsdS+ araD139 Δ(ara-leu)7697 Δ(lac)X74 galE15 galK16 rpsL mcrA mcrB1	Lab strain
TG-1	supE thi-1 Δ(lac-proAB), Δ(mcrB-hsdSM)5 [F´ traD36 proAB lacIqZΔ M15]	Lab strain
Xl-1 blue	recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac, (F' proAB, lacIq, lac z ΔM15, Tn10 (tetR))	Stratagene
pAK400	*lac* promoter, pUC origin, cm^R	(Krebber et al. 1997)
pTJ-PFP	pAK400, p_{lac} *sstorA-jun, sspelB-fos-pIII(198-406)*	Tullman-Ercek, unpublished
pP-GFPmut2-HPQ	pAK400, p_{lac} *sstorA-jun-gfpmut2(AGHPQ::157), sspelB-fos- pIII(198-406),*	Tullman-Ercek unpublished
pP-GFPrft-half	pAK400, p_{lac} *sstorA-jun-gfprft(1-157), sspelB-fos- pIII(198-406),*	This study
pP-GFPrft	pAK400, p_{lac} *sstorA-jun-gfprft, sspelB-fos-pIII(198-406)*	This study
pP-GFPrft-HPQ	pAK400, p_{lac} *sstorA-jun-gfprft(AGHPQ::157), sspelB-fos- pIII(198-406),*	This study
pP-GFPrft-peptide	pAK400, p_{lac} *sstorA-jun-gfprft(BglII-NNB_6::157), sspelB-fos- pIII(198-406),*	This study

Table 4.2 Oligonucleotides

Name	***Sequence***
Gfp-NheI	GCGGCGGCTAGCAGTAAAGGAGAAGAACTTTTCACTGGAGTTG
Gfp-SphI	CGCCGCGCATGCTTTGTATAGTTCATCCATGCCATGTG
Gfprev157HPQ	GCCCTGCGGATGACCCGCTTGTTTGTCTGTCATGATGTATAC
Gfpfor158HPQ	GCGGGTCATCCGCAGGGCAAGAATGGAATCAAAGCTAACTTCAA AATTAG
Gfprft157rev	CGCCGCGCATGCGACGTCAGATCTTTGTTTGTCTGTCATGATGTA TAC
Gfprft-158NNB	GCGGCGAGATCTNNBNNBNNBNNBNNBNNBAAGAATGGAATCA AAGCTAACTTCAAAATTAG

4.2.2 Library construction

The N-terminal region encoding amino acids 1-157 of GFP was amplified with the primer Gfp-NheI containing NheI cut-site for the 5' end and the primer Gfprft157rev containing the cut-sites BglII, AatII and SphI at the 3' end. The PCR product was then digested with NheI and SphI, and ligated into the plasmid pTJ-PFP (pAK400 containing *sstorA-jun, sspelB-fos-gIII(198-406)*, downstream of *jun*, giving rise to plasmid pP-GFPrft-half. The C-terminal half of GFP was amplified with the primers Gfprft-158NNB and the reverse primer Gfp-SphI with a SphI cutsite. Gfprft-158NNB contains a BglII restriction site at the 5' end followed by 6 random NNB codons (B can be C,G or T) before the sequence for the C-terminal half starts (codons for amino acids 158-168). The peptide library and the C-terminal half of GFPrft were ligated between the BglII site and the SphI site in the plasmid pP-GFPrft-half.

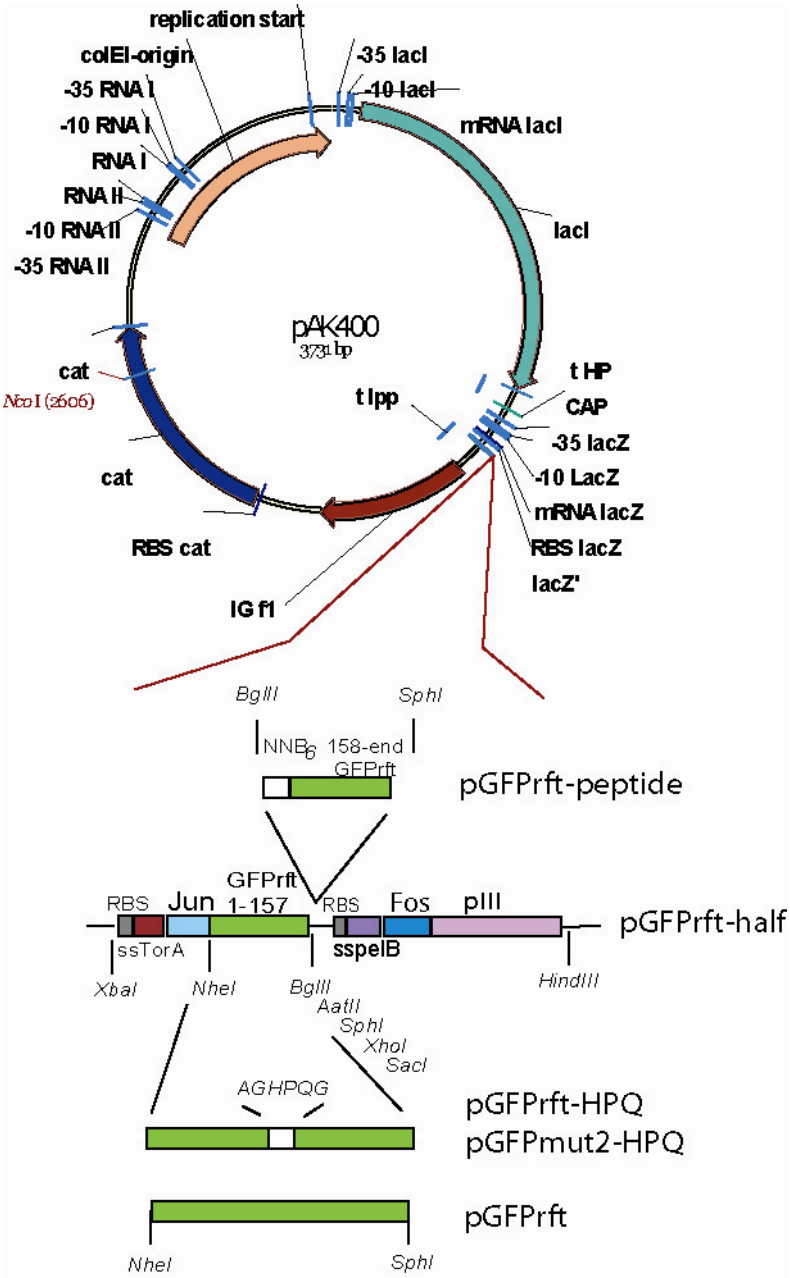


Figure 4.3. Map of phagemid and all constructs

4.2.3 Flow cytometry

DHF cells harboring the phagemid encoding the GFP peptide insertion variants were grown in 50 ml of LB media containing 1% glucose at 37˚C and 25 μg/ml chloramphenicol. Upon entry into the mid-logarithmic phase ($O.D._{600}$ 0.5-0.6), cells were transferred into induction medium containing LB with 0.1 mM IPTG and 25 μg/ml chloramphenicol. After 3.5 h, 10^8 cells were washed once in PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4) and sorted on a Becton Dickinson FACSAria flow cytometry instrument (excitation 488/emission 530 nm, fluorescence above a geometric mean of 200). Immediately after the first round of sorting, cells were sorted a second time, concentrated by filtration and then plated on LB plates containing 1% glucose and 25 μg/ml chloramphenicol.

4.2.4 Phage amplification

Cells containing the plasmid pAK400 *torA-jun-gfp, pelB-fos-gIII,* with or without peptide insertion into GFP, were grown in 200 ml 2xYT media containg 1%glucose and 25 μg/ml chloramphenicol at 37˚C following inoculation to an $O.D._{600}$ =0.1. Upon entry into the mid-logarithmic phase, cells were infected with M13K07 helper phage in a 1:20 ratio and incubated at 37˚C without shaking for 1 h. Phage remaining in the supernatant was eliminated by centrifugation and re-suspension of the cells in fresh 2xYT medium containing the 25 μg/ml chloramphenicol, 0.1 mM IPTG to induce expression and 25 μg/ml kanamycin to select for cells containing helper phage. After overnight expression and phage propagation, phage particles were precipitated using 1/5 volume of 20% PEG, 2.5 M NaCl, typically overnight or for 4 h at 4˚C. Precipitated phage was resuspended in

2 ml PBS with 15% glycerol and the titer determined. Typically, titers of around 10^{12} colony forming units/ml were achieved. Once the optimum conditions for GFP display on the phage were established, cells were typically grown at 30˚C for phage assembly.

4.2.5 Phage panning using immunotubes

4.2.5.1 Simple coating, simple panning

Immunotubes (Nunc) were coated with the 100 μg/ml antigen while rotating overnight at 4˚C. All following steps were performed at room temperature. Immunotubes were blocked with 2% milk in PBS for 2 h while rotating. After 2 wash steps with PBS, about 5 x 10^{11} phage particles (based on cfus) were resuspended in 2% milk PBS, and incubated in the immunotubes for 1.5 h while rotating. After another 30 min incubation in which the tubes were not rotated, the immunotubes were washed 10 times with PBST (PBS, 0.1% tween) and then 10 times with PBS to eliminate weakly bound phage particles. For each subsequent round, 5 washing steps for both PBS and PBST washes were added. Phage was eluted with 500 μl of 100 mM triethyl amine by rotating for 30 min. Phage elution was immediately neutralized with 1 ml of 1 M Tris pH 7.4. 700 μl of the neutralized, eluted phage solution was used to infect 6 ml of DHF cells in the mid-logarithmic growth phase. 4 ml of cells were incubated in the immunotubes. Infection was performed at 37˚C without shaking for 1 h. 10 μl of infected cells were used for serial dilution to determine the total cfus obtained by panning. The remaining cell suspension was concentrated by centrifugation and plated on LB plates with 1% glucose and 25 μg/ml chloramphenicol.

4.2.5.2 Blocking and preabsorption of immunotubes

For pre-absoprtion, phage particles were incubated in immunotubes that were coated with 100 μg/ml glucose oxidase before they were transferred to the tubes coated with the antigen of interest.

Typically, 5 x 10^{11} cfu of phage displaying the GFPrft peptide library were resuspended in PBS containing 2% milk and added to immunotubes which had been coated with 100 μg/ml glucose oxidase overnight and blocked for 2 h in the same buffer at room temperature. Immunotubes coated with the ligand of interest (100 μg/ml lysozyme or transferrin, respectively) were blocked for 2 h with 2% BSA in PBS instead of milk to avoid selection of binders to the blocking agent. After 1 h of rotating in the pre-absorption tube, the phage solution was transferred into the actual selection tube. The panning procedure was performed as described under section 4.2.5.1.

4.2.3 Phagemid panning using magnetic streptavidin beads

4.2.3.1 Pre-absorption step

80 μl of the Streptavidin dynabeads stock solution M-280 (Invitrogen) containing around 6 x 10^7 beads were were washed twice with PBS and resuspended in PBS. Beads were then blocked for 1 h while rotating. For each step, we altered the blocking agent. For odd numbered rounds of panning, beads were blocked with 2% milk whereas the phage was resuspended in 2% BSA in PBS. In even-numbered rounds, beads were blocked with 2% BSA in PBS and phage was resuspended in 2% milk PBS. Typically, a

total number of at least 10^{10} phage particles (cfus) were used as a starting concentration. Phage particles were incubated with blocked beads for 1 h while rotating at room temperature. Phage was then used for the following panning procedure.

4.2.3.2 Coating beads with transferrin

20 µg/ml of biotinylated transferrin (Invitrogen, molecular probes) which contains 5 biotin molecules per transferrin (as indicated by the manufacturer), was added to 6 x 10^7 M-280 beads (pH 7.4). After 30 min incubation of the beads with biotinylated transferrin while rotating, beads were washed once and blocked for 1 h with either BSA or milk as described under the preincubation step. After 3 wash steps using PBS, phage particles from the pre-absorption step were added to the transferring-coated beads and incubated for a specific time (see below).

4.2.3.3 Panning procedure

Unless noted differently, all the following steps were performed at room temperature while rotating. Phage particles recovered from the pre-absorption step were transferred to the transferrin-coated beads and incubated for 1 h; beads were then washed 3 times with each PBST and PBS to eliminate weak binders. Bound phage particles were eluted in 2 steps. For the first elution, beads were incubated for 10 min with 500 ul of 40 mM triethyl amine; as a second elution step, we used 100 mM triethyl amine (TEA). Eluted phage in TEA was immediately neutralized by suspension in 1 ml of 1 M Tris pH 7.4 buffer. The two different elutes were incubated independently with DHF cells in mid-logarithmic growth phase and plated. For the second round of panning, recovered obtained phagemids was amplified in phage as described above. For the second round of

panning, blocking was performed as described above, and two extra wash steps for each PBST and PBS were applied to eliminate weak binding phage particles. Further, the beads were rotated for an extra 2 h in PBS before elution. Elution was performed as described above. After the 1st round, we only amplified phagemid of cells recovered form the 2nd elution. However, all recovered phagmid was amplified in phage after the 2nd round of panning before proceeding to the 3rd round. Purified phage particles were exposed to the 3rd round in which the phage incubation steps with transferrin-displaying beads was shortened to 45 min and the washing conditions were intensified: The last wash step using PBST was extended to 15 min rotating in PBST, before continuing with 5 PBS washes. The 4th round was essentially performed as the 3rd round of panning, except that the washing step in which the phage particles were rotated in PBST was extended to 30 min.

4.2.4 Single clones analysis for fluorescence

96-well plate overnight cultures of cells grown in 2xYT media containing 1% glucose and 25 μg/ml chloramphenicol were subcultured into the same media. After growth to mid-logarthmic growth phase (around 2 h), cells were centrifuged and resuspended in fresh 2xYT containing 0.1 mM IPTG and 25 μg/ml chloramphenicol. After 4 h, cells were washed in PBS and fluorescence was measured using a BioTek plate reader (exitation 485nm, emission: 528 nm).

4.2.5 Phage ELISA analysis

Phage ELISA analysis of individual clones was performed to identify GFP variants that bound target selectively. Cells were grown in 96-well plates in 200 μl 2xYT

containing 1% glucose and 25 μg/ml chloramphenicol overnight. 20 μl of the overnight was used to inoculate 200 μl 2xYT containing 1% glucose and 25 μg/ml chloramphenicol. After growth to mid-logarithmic growth phase, cells were infected with M13KO7 helper phage in a 1:20 ratio, cell count was estimated on their $O.D._{600}$. After 1 h of infection at 37˚C without shaking, cells were centrifuged at 4400 xg and resuspended in fresh 2xYT medium with 0.1 mM IPTG, 25 μg/ml kanamycin and 25 μg/ml chloramphenicol. Phage was produced overnight at 30˚C. 25 μl of the culture supernatant was used for ELISA analysis. As control for unspecific binding either laccase, BSA, glucose oxidase or streptavidin was used.

4.2.6 ELISA

Reacti-Bind™ White Opaque high-binding plates (PIERCE) were coated overnight with 100 μg/ml of the ligand (transferrin lysozyme or streptavidin), and either glucose oxidase, BSA, streptavidin or laccase as a negative control. After 2 washing steps with PBS, plates were blocked with 2% milk PBS for 2 h, washed once with PBST and 3 times with PBS. 25 μl of culture supernatant were added to 25 μl of 2% milk in PBS, so that each culture was tested for binding to the specific antigen and a negative control such as glucose oxidase, laccase or streptavidin. After 1.5 h, plates were washed one time with PBST and 3 times with PBS. Anti M13 antibody (Pharamcia) was applied for 1 h in a 1:5000 dilution in 2% milk PBS. Unbound antibody was rinsed off by washing 3 times with PBST and PBS each. For titration ELISA analysis, typically six serial dilutions were performed.

4.3 RESULTS AND DISCUSSION

4.3.1 Generation of a fluorescent GFP peptide insertion library

4.3.1.1 Introduction

The engineering of ligand binding affinities in GFP has been approached by several research groups (Zeytun et al. 2003; 2004; Paschke et al. 2007). However, so far it has not been possible to insert a diverse repertoire of peptide sequences into fluorescent GFP. GFP variants with loop insertions were found to exhibit low solubility and decreased fluorescence (Zeytun et al. 2003; 2004; Paschke et al. 2007). Further, since GFP cannot be exported via the Sec pathway in *E. coli,* it is not compatible with the conventional phage display system (Feilmeier et al. 2000; Zeytun et al. 2004). Surface display on yeast could be potentially employed since secreted GFP can assume its fluorescent conformation in this organism (Huang and Shusta 2005); however the yeast display system is limited by the size of variant libraries that can be screened since transformation efficiencies are generally lower than for *E. coli*.

Recently, a Tat-based phage display system was developed by the Georgiou lab and others (Paschke, 2005; Tullman-Ercek unpublished results). In our Tat-based phage display system, the protein of interest is expressed as a C-terminal fusion to Jun and tagged with a Tat signal peptide. After export through the Tat translocon, it associates with the Fos-pIII fusion construct which was exported via the Sec system using a PelB signal peptide. The heterodimer is then assembled onto the phage particles. In this manner, it is possible to generate phage particles displaying fluorescent w.t. GFP or GFP mutants. Here, we describe the optimization of phage displaying fluorescent GFP and the generation of a large pool of fluorescence GFP variants containing a sixmer peptide insertion. However, so far it has not been possible to isolate fluorescent GFP variants

that bind to ligand, indicating perhaps structural limitations in the repertoire of sequences that can be inserted within GFP and still maintain its structural integrity and ability to form the chromophore.

Initially, we tested the GFP FACS-optimized version mut2 (Cormack et al. 1996) for its tolerance to a sixmer insertion (AGHPQG) at position 157 (Figure 4.1). The fusions ssTorA-Jun-GFP and ssPelB-Fos-pIII were encoded in a bi-cistronic operon in the phagemid pAK400 and the peptide was inserted at position 157 by site-directed mutagensis. We detected low fluorescence, sensitivity to higher temperatures (37˚C) and low phage titers derived from cells expressing GFP with the insertion (data not shown and Tullman-Ercek unpublished results). We therefore decided to utilize a more stable variant of GFP that affords more efficient folding at 37˚C and better solubility. Further, the GFP of choice, GFPrft, fused to the SsrA tag is relatively resistant to the ClpXP degradation (Karzai et al. 2000) (data not shown). Briefly, GFPrft14, the parent gene of GFPrft were generated by iterative rounds of error-prone PCR/shuffling and screening for maximum cell fluorescence (excitation at 360 nm) at 37˚C. The GFPrft14 variant contains several mutations that have been isolated in screens for brighter and more stable variants such as the cycle 3 variant (Crameri et al. 1996; Siemering et al. 1996) as well as the superfolder variant (Pedelacq et al. 2006). The mutations in GFPrft14 were Q80R, E124R, E124V, A154T and V163A. The variant GFPrft14 was further exposed to site-directed mutagenesis to change F64L and Ser65 into a threonine residue which causes more efficient excitation at 488 nm optimizing it for the use of flow cytometry (Patterson et al. 1997). The F64L/S65T gave rise to GFPrft (Yasuaki Kawarasaki, unpublished data) which we used in the present study.

In their characterizations of the superfolder GFP variant, Pedelacq and coworkers compared the solubility and fluorescence of several circularly permutated GFP variants

of both the w.t. protein and the engineered superfolder GFP variant (Pedelacq et al. 2006). Based on their results, it seemed that circular permutation at position 157 allows the highest fluorescence and solubility for both w.t. GFP and of the superfolder variant. Similar observations were made by Paramban and coworkers (Paramban et al. 2004). Therefore, we decided to insert peptides between the amino acids 157 and 158.

4.3.1.2 Fluorescence analyses

Cells were grown to mid-logarithmic growth phase at 30˚C and induced with 0.1 mM IPTG. After a 3 h growth period, uninduced and induced cells were compared via FACS on a Becton Dickinson FACSSort instrument. At least at 30˚C, we were able to see fluorescence and efficient repression of the protein construct in the absence of the inducer (Figure 4.4). Efficient repression is important to avoid any growth disadvantage of clones expressing truncated polypeptides that may otherwise overtake the library.

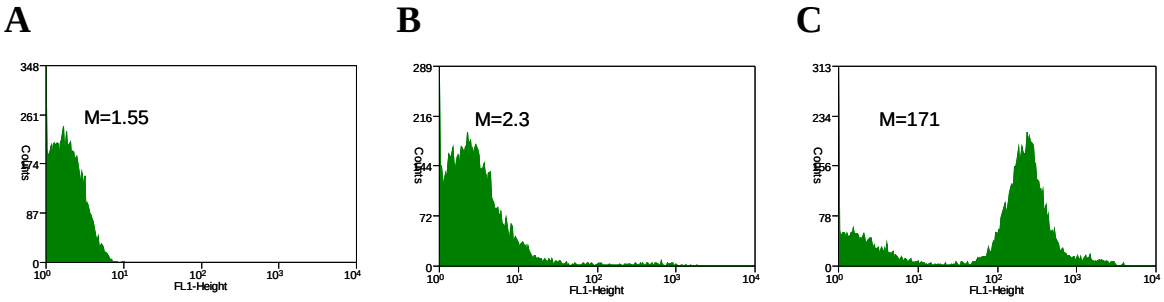


Figure 4.4. Fluorescence in TG-1 cells after 3 h induction period at 30˚C. (A) TG-1 cells, (B) TG-1 cells with uninduced pAk400 *sstor-jun-gfpmut2(HPQ), pelB-fos-gIII*. (C) TG-1 cells harboring pAk400 *sstor-jun-gfpmut2(HPQ), pelB-fos-gIII* were induced with 0.1 mM IPTG.

4.3.1.3 Optimizing expression and fluorescence of GFP containing peptide insertions

Since the GFP Mut2 with a peptide insertion showed low levels of fluorescence (Tullman-Ercek, Yasuaki Kawarasaki, and unpublished observation) and resulted in low phage titers, we proceeded to use the temperature stable GFPrft variant. In addition we switched the host cell to *E.coli* DH5α which has a slower doubling time and a deletion of the *recA* gene the reduces homologous recombination. Since DH5α does not contain a F-pilus which is essential for filamentous phage infection, we crossed DH5α with XL-1 blue (*recA1, endA1, gyrA96, thi, hsdR17, supE44, relA1, lac,* [F^+ *proAB, lacIq, lacZ ΔM15, Tn(*tetR*)*]) giving rise to *E. coli* strain DHF. DHF expressing GFP mut2 at 37˚C gave substantially higher fluorescence relative to TG-1 cells, but the GFPrft variant proved to be more tolerant to a peptide insertion than GFP mut2 (Figure 4.5).

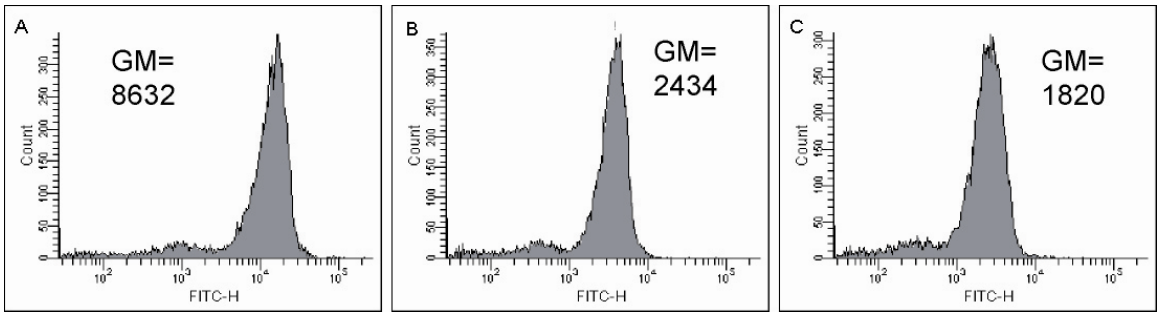


Figure 4.5. Fluorescence in DHF cells 3 h after induction. Induced cells harboring the plamids (A)pP-GFPrft, (B)pP-GFPrft-HPQ and (C) pP-GFPmut2-HPQ. Whereas (B) and (C) are expressing a GFP variant containing a 6aa insertion (AGHPQG) at position 157.

We proceeded to generate a peptide insertion library by first engineering a BglII site at the codon for aa 157 of GFPrft. A DNA fragment encoding the C-terminal half of GFPrft with six NNB codons (B includes all nucleotides but A) at its N-terminus and having BglII and SphI sites at the 5' nd 3' ends was generated by overlap extension PCR.

The PCR product was then ligated within the corresponding sites of the plasmid pP-GFPrft-half. For transformation of the library, we utilized the *E. coli* strain MC1061 which has proven to have high transformation efficiency for ligated plasmid constructs due to the deletion of the endonucleases encoded by mcrA/B and hsdR. About 10^7 individual clones were obtained and the diversity of the 6-mer insertion was verified by sequencing 9 clones selected at random (Figure 4.6). One *gfp* gene contained an extra codon insertion next to the 6 random amino acid codons (not shown). Around 30% stop codon are anticipated when using the NNB as a random insertion (Ostermeier 2003).

Figure. 4.6. Sequence alignment of peptide insertions from random library variants

4.3.1.4 Enrichment for fluorescent GFP containing peptide insertions

To enrich cells expressing fluorescent GFP variants, cells containing the GFPrft peptide library were screened by flow cytometry (FACSAria). The fluorescence histogram (ex. 488 nm/em. 530 nm) indicated that 18% of the population was fluorescent after a 3.5 h induction period at 37˚C. The sorting gate was set so that cells with fluorescence higher than 200 a.u. (a.u.=arbitrary units) were collected. In addition, a forward scatter (FSC) threshold was employed to gate on cells and eliminate debris from the sort solution. The forward scatter is directly proportional to the size of cells examined by the laser of the flow cytometer. The FSC detector is placed in line with the laser path

and the diffracted light is collected. The bigger the cells the more light is scattered and the higher the FSC signal. Immediately after the first round of sorting, cells were resorted and then plated on LB media with 1% glucose and 25 µg/ml chloramphenicol. 2 x 10^6 colonies were obtained reflecting approximately the expected percentage of fluorescent variants within the library pool. The overall fluorescence of the selected population was around 100 fold above background as detected in a FACSAria instrument (Figure 4.7.A). The diversity of the library was confirmed by sequencing 10 random clones (Figure 4.7.B).

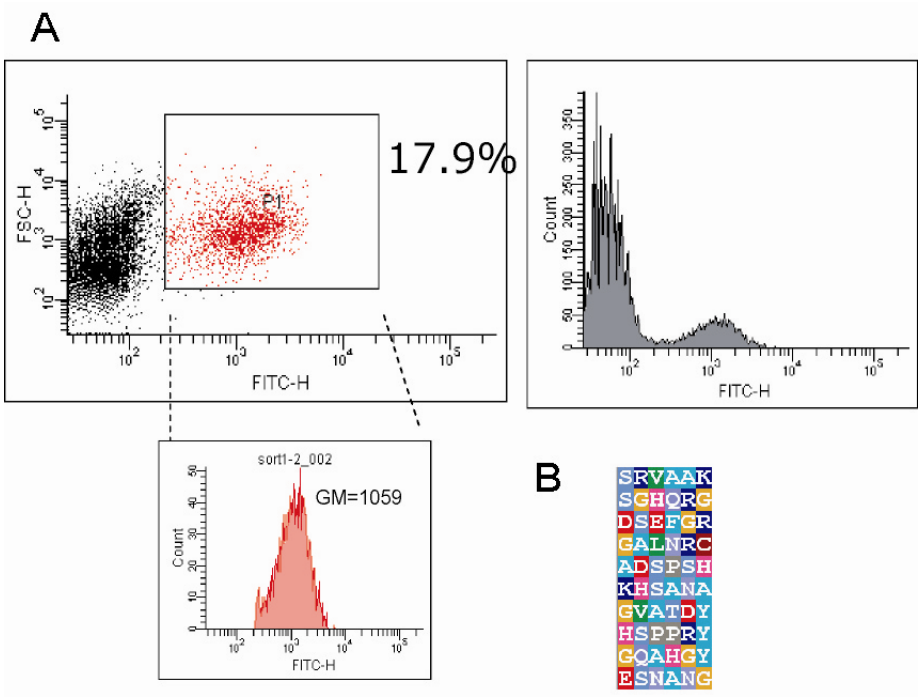


Figure 4.7. (A) Fluorescence histograms of the GFP peptide library, (B) sequence alignment of 10 random clones after 2 library sorts.

4.3.2 Optimization of GFP display on M13 phage

The enrichment of fluorescent variants of the peptide insertion library was performed with cells grown at 37˚C. However, the conditions for optimum display of fluorescent GFPrft on M13 filamentous phage needed to be determined. Cells harboring either pP-GFPrft or pP-GFPrft-HPQ (i.e. containing the 6 aa insertion AGHPQG) were grown with or without induction at 30˚C or 37˚C and superinfected with helper phage to allow phage propagation.

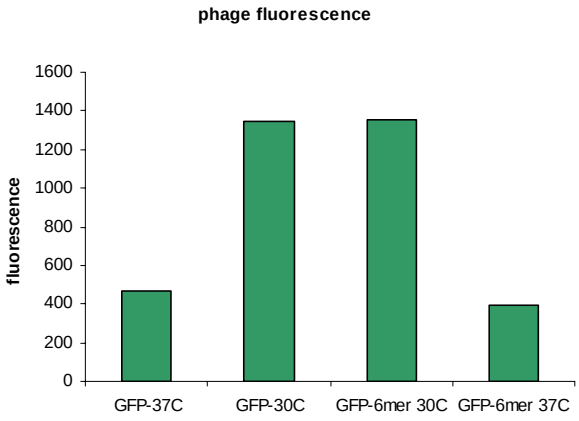


Figure 4.8. Fluorescence of M13 phage displaying GFPrft. Different growth conditions were tested to optimize display, temperatures are indicated. Phage solutions were normalized to 10^{11} cfu/ml, 200 µl of this solution was analyzed in black 96-well plates (Nunc) using a BioTek plate reader (ex. 485 nm/em. 528 nm).

In cells that were not induced, GFP expression could not be detected (data not shown) and therefore we did not analyze the fluorescence of phage particles amplified without an inducer. The phage solution was normalized to a titer of 10^{11} colony forming units (cfu)/ml and 200 µl of purified phage was analyzed for fluorescence (using a BioTeck ELISA reader (excitation 485 nm, emission 528 nm) in black plates with

transparent bottom (COSTAR®)) (Figure 4.8). Fluorescence of the phage was dependent on the temperature at which the phage assembly had occurred. Optimal display of GFPrft was achieved when the phage particles were assembled at 30˚C. Insertion of the 6 aa peptide into GFPrf did not reduce the fluorescence of the phage particles

The fluorescence data shown above might be due to GFP protein that is precipitated together with phage rather than intact phage particles. To rule out this possibility we examined whether the phage displayed GFP that could recognized by anti-GFP antibodies in ELISA assays.

Phage particles, which were assembled in cells growing at 30˚C, and either GFPrft or pGFPrft-HPQ, were coated onto immunoplates by serial dilutions. After blocking the wells were reacted with monoclonal anti-GFP antibody and then a secondary anti-mouse horse radish peroxidase (HRP) conjugate was applied

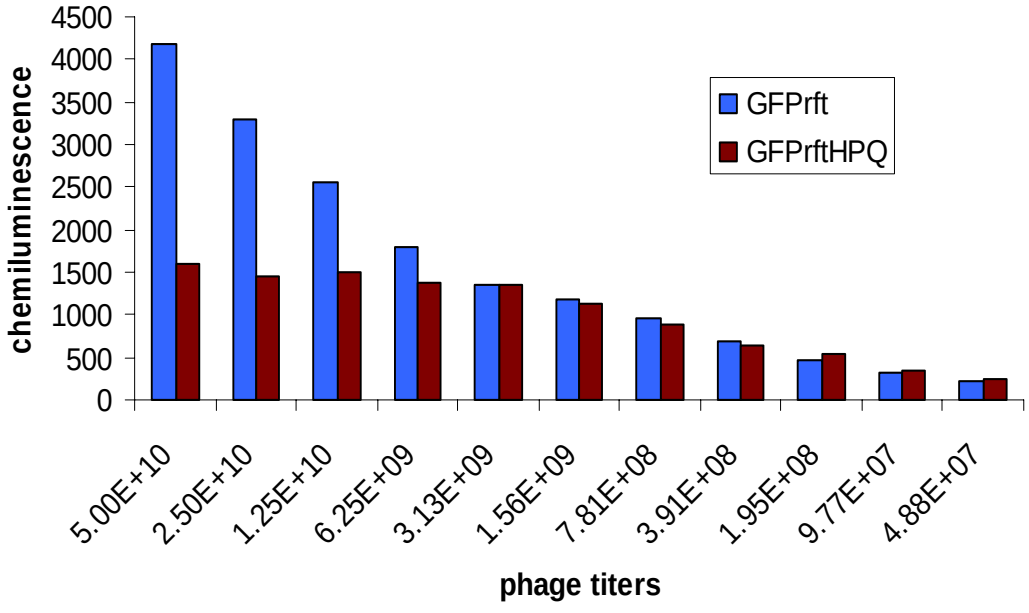


Figure 4.9. ELISA analysis on plates coated in serial dilutions of phage displaying either GFPrft or GFPrft-HPQ. Opaque plates were coated with serial dilutions of GFP displaying phage. GFP was detected with a monoclonal antibody anti-GFP antibody which was in turn recognized by a anti-mouse HRP conjugate. HRP activity was monitored by chemiluminescence generated by the HRP-catalyzed oxidation of Supersignal® Western Pico chemiluminescence substrate (PIERCE). Signals were recorded by a BioTek plate reader using a 460 nm filter.

The ELISA data indicated that GFPrft is displayed on the phage particle (Figure 4.9). Surprisingly, the anti-GFP antibody used to detect GFPrft with a peptide insertion anchored to the phage reached saturation before the w.t. GFPrft.

4.3.3 Selection for ligand-binding GFP via phage panning

Three different proteins were chosen as possible ligands for panning experiments: streptavidin, lysozyme and transferrin. Streptavidin is a 60 kDa tetrameric protein isolated from *Streptomyces avidinii* with high affinity for biotin (10^{-15} M affinity). The 14.4 kDa lysozyme hydrolyzes 1,4-beta-linkages between N-acetylmuramic acid and N-acetyl-D-glucosamine residues which are found in bacterial cell walls. Human transferrin is a 80 kDs glycoprotein that binds to its receptor located in capillary endothelial cells and delivers iron from the blood to the brain. Through its interaction with the receptor, transferrin crosses the blood-brain barrier via transcytosis. Monoclonal antibodies developed to the transferrin receptor are currently used as Trojan horses to deliver nonlipophilic drugs to the brain tissue (Friden et al. 1991).

4.3.3.1 Enrichment for antigen specific binders via immunotubes

For phage display, ligands can be presented either in soluble form or immobilized on beads or immunotubes. Panning steps in which the antigen is presented in a solution are considered to be "stringent" and therefore used for the isolation of higher affinity binders (Barbas III et al. 2000). In immunotubes, ligands are immobilized and thus multivalent binding that results in higher avidity interaction between the phage and the immobilized ligand is possible. However, not every antigen binds efficiently to the plastic surface of the tubes. Initially, panning relied on immobilization of the ligands to the surface of immunotubes. Coating and panning steps are described in detail under the method section. The number of colony forming units after each panning step is listed Table 4.). Each consecutive round of panning showed a significant enrichment resulting in 10 fold or even a 100 fold increase in the number of colony forming units for streptavidin, lysozyme or transferrin.

Table 4.3. Enrichment of phage particles binding to the protein of interest. Numbers correspond to colony forming units (cfus).

Round of panning	***anti-streptavidin***	***anti-lysozyme***	***anti-transferrin***
1^{st} round	8×10^3	3×10^3	10^4
2^{nd} round	10^5	10^6	10^6
3^{rd} round	1.3×10^8	5×10^7	N/D

To verify that enrichment was due to the binding of phage onto streptavidin, lysozyme or transferrin, respectively, polyclonal phage ELISAs were performed

4.3.3.2 Enrichment of streptavidin-binding GFP

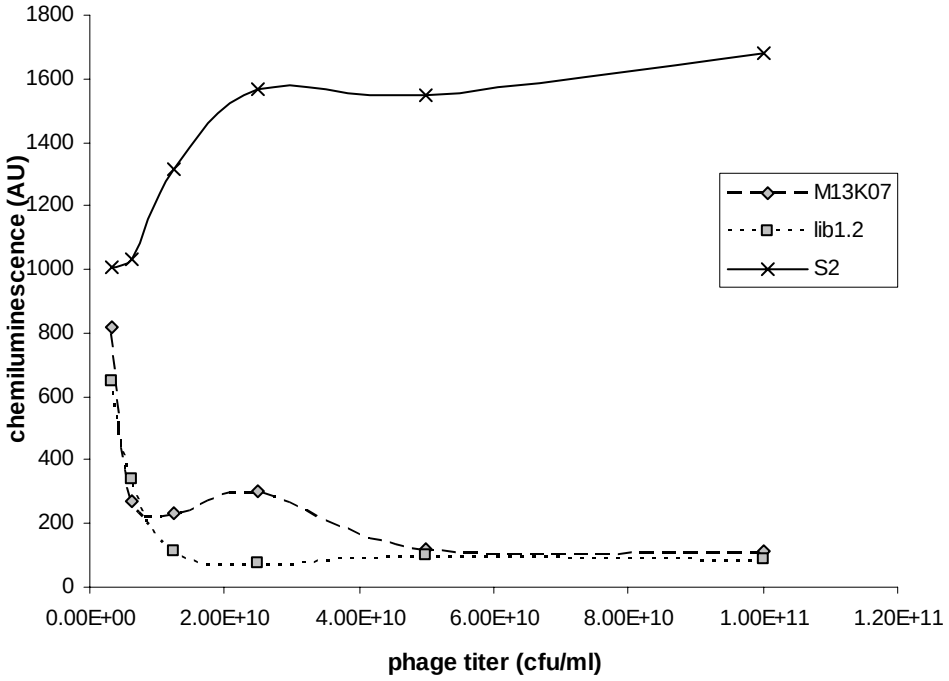


Figure 4.10. Polyclonal phage ELISA against streptavidin. S2 phage purified after the 2nd round of panning. Lib1.2 represents isolated fluorescent GFP variants after 2 rounds of sorting using FACS. M13K07 is solely helper phage which is not displaying any foreign protein.

Binding of phage displaying GFPrft variants to streptavidin was observed after the second round of panning (Figure 4.10). We further performed a third round of panning, which resulted in a 1000 fold increase in colony forming units (Table 4.3). Individual clones from the second and third round of panning were selected at random. 72 clones were analyzed for fluorescence in a 96-well plate format (see method section). 25 clones which gave a fluorescence signal were further analyzed for ligand binding. Cells containing the respective phagemids were subcultured and infected with helper

phage. After overnight propagation of the phage under conditions in which pP-GFPrft-peptide variants are expressed, the supernatant was collected, and binding to streptavidin was detected by ELISA. However, only one clone was shown to exhibit binding to streptavidin in a concentration dependent manner (Figure 4.11).

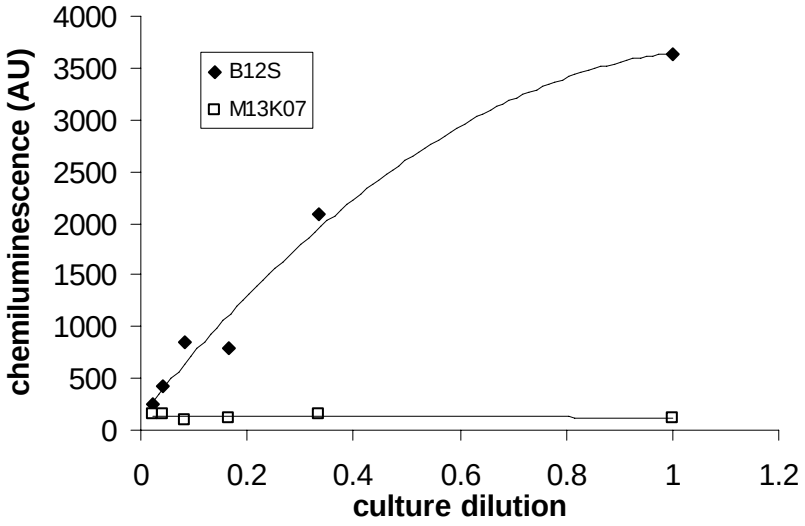


Figure 4.11. ELISA analysis of culture supernatant expressing the GFPrft isolate B12S Empty phage (M13K07) was used as a control.

Upon concentration of the phage by precipitation with PEG, clone B12S was found to exhibit weak binding to streptavidin that was only about 2-2.5 times higher than the binding of the phage to BSA (Figure 4.12).

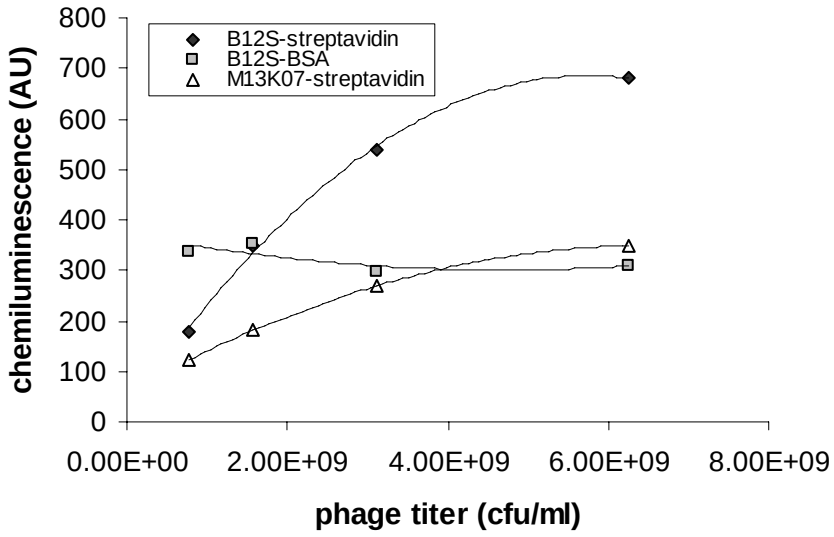


Figure 4.12. Phage ELSIA analysis using precipitated phage expressing B12S

Considering the high amount of phage (>10^9 cfu) needed to detect binding of the B12S phage onto streptavidin, it appears that the binding affinity is very low. The difference between the signal obtained with culture supernatants and with purified phage may be due to the conditions used for phage precipitation. It is possible that the high salt concentration (a total of 0.5 M NaCl) used to precipitate the phage particles disrupted the interaction of Jun and Fos. This could result in lower amounts of GFP anchored to the phage which would explain the difference between ELISA signals obtained from supernatant of the cultures versus purified phage particles. Unfortunately, FACS analysis of cells expressing the B12S variant of GFPfrt revealed that this protein exhibits very low fluorescence, only slightly above background (2 fold) compared to 240 fold about background for the GFPrft-HPQ protein (data not shown). This observation contradicted our earlier observation when using 96-well plate fluorescence analysis. FACS analysis allows the detection of fluorescence per cell whereas the 96-well plate analysis only

measures the overall fluorescence. Cells expressing this variant grew faster which increased the overall fluorescence when compared to the other GFPrft variants.

In summary the above analysis indicates that panning for strepatividin binders resulted in the isolation of a weakly binding clone in which the peptide insertion interfered with the chromophore maturation of GFP.

4.3.3.3 Panning against lysozyme

3 rounds of panning against lysozyme coated on immunotubes were performed. Polyclonal phage ELISA showed signals only slightly above background.

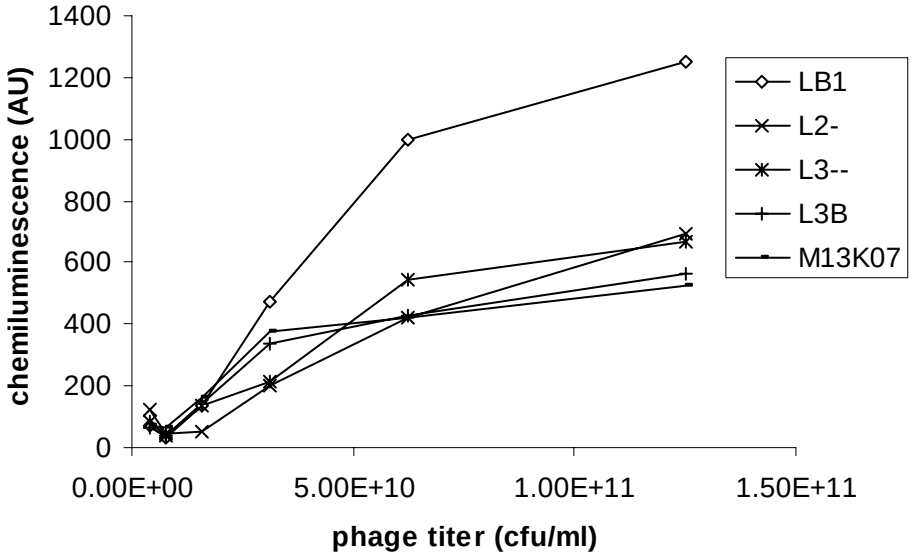


Figure 4.13. Polyclonal phage ELISA analyses after different rounds of panning against lysozyme. LB1 is the first round of panning, whereas LB3 was the third round of panning. L2- and L3--were recovered after panning with pre-incubation steps (see section 4.3.3.3.1).

The ELISA signal decreased after the first round of panning, indicating that non-binding mutants or possibly truncated versions of GFPrft may have outgrown the binding variants during the phage amplification step. Nonetheless, we proceeded to analyze

several clones isolated from different rounds. Screening of 25 fluorescent clones (out of 60 tested) led to the identification of one clone, 5A9L which showed a concentration dependent binding to lysozyme when analyzed by ELISA of the culture supernatant (Figure 4.14).

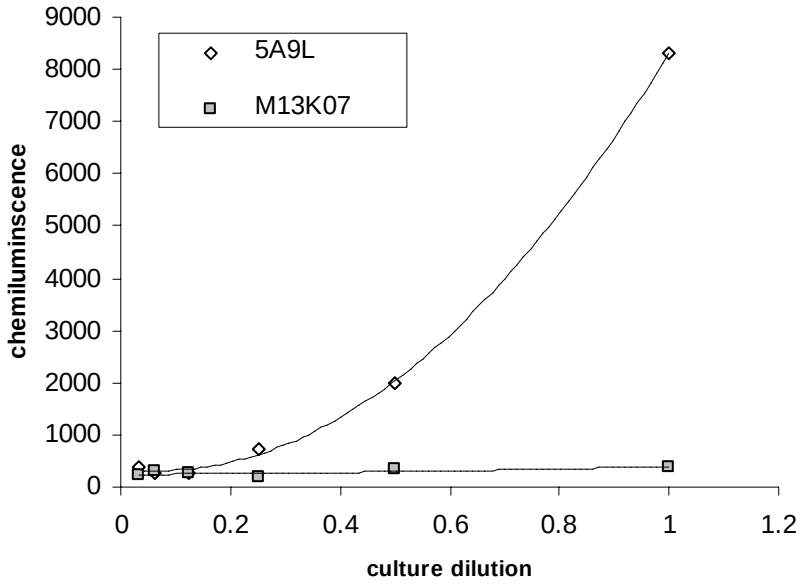


Figure 4.14. ELISA assays of culture supernatants expressing the GFPrft isolate 5A9L. Empty M13K07 was used as a negative control.

However, similar to the streptavidin binder B12S discussed above, the ELISA signal was substantially lower when phage was precipitated with PEG for purification. Specifically, the ELISA signal was barely above background binding when compared to w.t. helper phage (M13K07) (Figure 4.15). We further detected non-specific binding to BSA and glucose oxidase (data not shown).

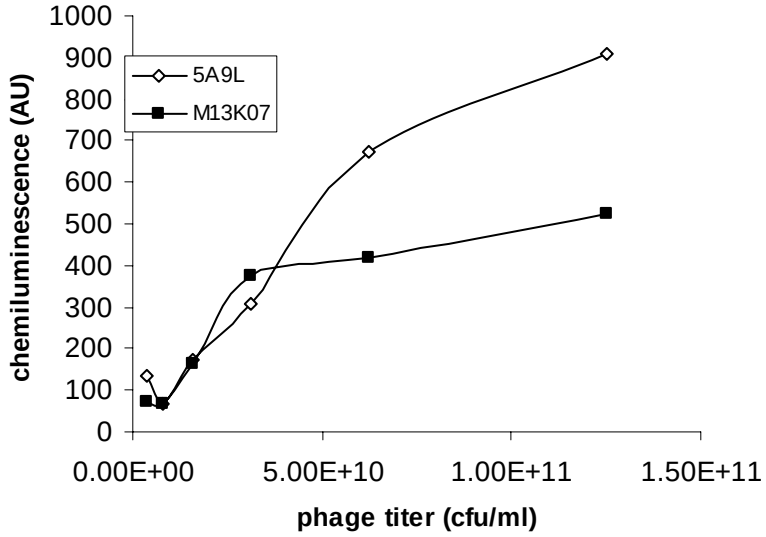


Figure 4.15. ELISA analysis using PEG precipitated phage expressing 5A9L and empty phage (M13K07) as a control.

4.3.3.3.1 Addition of a pre-absorption step

Since the panning procedure used above against streptavidin or lysozyme enriched only a couple of low affinity binders which seem to be non-specific, we decided to alter the panning procedure. Clones recovered after the first round of panning against lysozyme were exposed to a pre-absorption step using immunotubes coated with an unrelated protein (glucose oxidase). Phage particles displaying the GFPrft variants were exposed for 1 h to glucose oxidase-coated tubes before binding to lysozyme coated tubes was allowed. Phage elution resulted in 3.5 x 10^5 cfu. The decrease in colonies compared to the second round of panning without preabsorption (Table 4.3), indicated that non-specific binders were likely reduced from the pool of phage. We proceeded to perform a third round of panning using this protocol, which resulted in an output of 3.6 x 10^7 cfu.

96 clones were selected at random and analyzed by ELISA. We identified 5 clones that bound with higher affinity to lysozyme than to streptavidin (Figure 4.16).

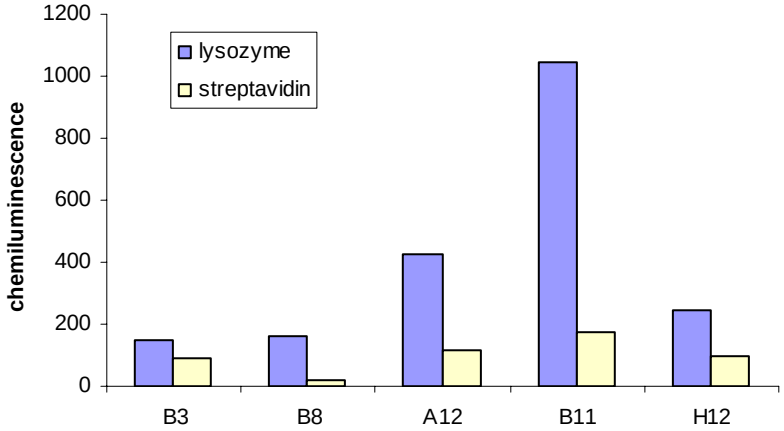


Figure 4.16. Lysozyme binders after 3 rounds of panning in which the last 2 rounds were accompanied by a pre-absorption step. ELISA singals were obtained from culture supernatant

However, as seen before, the result could not be reproduced when the phage particles displaying the GFPrft variants were concentrated by PEG precipitation.

4.3.3.4 Isolation of GFPrft variants binding to transferrin

As the polyclonal phage ELSIA indicated, no specific binding clones were detected after the first 3 rounds of panning against transferrin. However an elevated polyclonal phage ELISA signal was seen with phage particles obtained after the second round of panning (Fig.4.17).

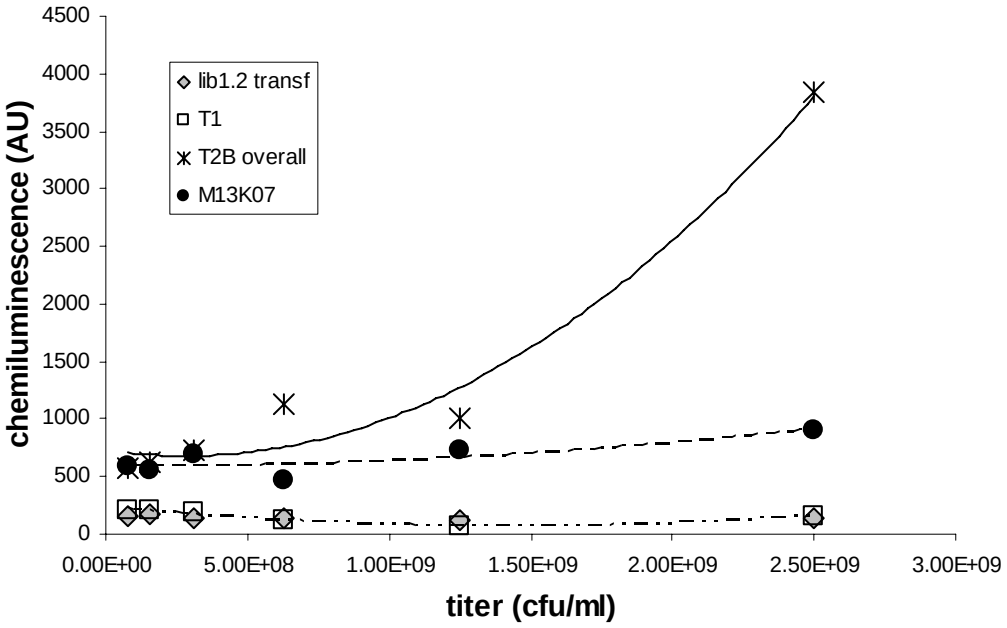


Figure 4.17. Polyclonal phage ELISA against transferrin using purified phage particles after the second round of panning.

We performed monoclonal phage ELISAs on phage isolated after the 2nd and 3rd round of panning, but even though 200 different variants were analyzed, we could not identify any binders that gave a signal above background. We therefore used the population of phage particles after the second round of panning which showed binding to transferrin and performed a third round of panning using the preincubation protocol with

glucose oxidase-coated tubes. Surprisingly, only 3 x 10^3 colonies were obtained after elution and infection of fresh cells. We analyzed culture supernatants from 200 clones for binding to transferrin with laccase or BSA as negative controls. We identified 10 clones that gave slightly higher ELISA signal for transferrin than to the negative control (Figure 4.16). Differences in the ELISA signals for the negative control are due to different phage concentrations in the supernatant of the cultures (Figure 4.18).

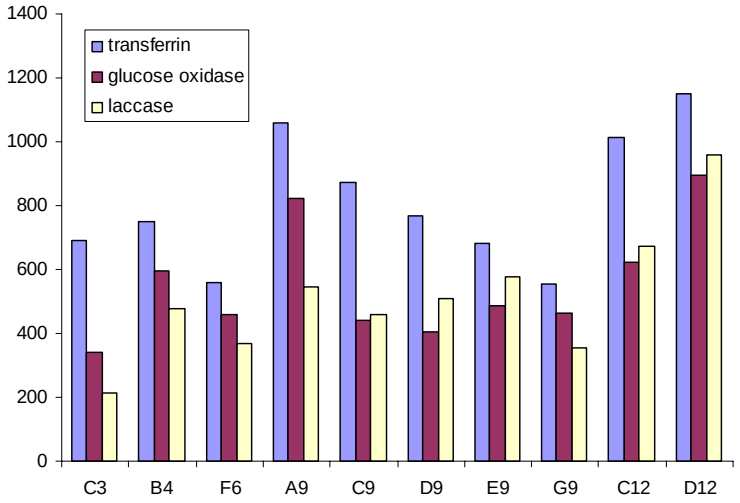


Figure 4.18. Phage ELISA on supernatant of infected cells grown in 96-well plates. Clones were picked after 3 rounds of panning in which the 3rd round was accompanied by a pre-absorption step using immuntubes coated with glucose oxidase. 10 out of 192 tested clones showed higher binding affinity towards transferrin than towards the negative control.

All 10 clones which showed higher affinity towards transferrin than glucose oxides or laccase (Figure 4.16) were grown in 4 ml cultures, infected with helper phage, and precipitated with PEG/NaCl. ELISA analysis using varying phage concentrations were performed using transferrin and glucose oxidase as a model unrelated protein.

However, all isolates showed higher affinity towards glucose oxidase than transferrin (data not shown).

4.3.4 Magnetic bead-based selection for transferrin binding GFP

To circumvent the possibility that transferrin did not bind well to the surface of the immunotubes, thus resulting in the negative results discussed above, we used protein immobilized on magnetic beads. M-280 Dynabeads (Invitrogen) coated with streptavidin were used to pull-down biotinylated transferrin (Invitrogen), which ensures the immobilization of transferrin on the beads. However, it also allows binding to the bead surface to biotin or to streptavidin. To avoid enrichment to the latter, pre-absorption steps were included using blocked beads without transferrin. To further minimize the possibility of an enrichment of binders to the blocking agent, we altered the blocking solution for each round of panning (see in detail under section 4.3).

4.3.4.1 Panning rounds using the Dynabead system

As mentioned above a pre-absorption step was performed and the blocking agent was altered in each round. The cfus obtained after each round of panning using the dynabead system are summarized in Table 4.4. The increase of the cfus for each round of panning indicates enrichment. We further differentiated between phage isolated after the first and second elution step from phage bound to the bead system.

Table 4.6. Overview of enrichment for cfus using bead approach

Round of panning	*M-280 (cfu)*	*M-280 (cfu), 2nd elute (from first round)*	*M-280 (cfu), beads*
1st round	10^4	6×10^3	2×10^3
2nd round	2.97×10^6	$1.5\ 10^6$	10^5
3rd round	N/D	1.12×10^7	7×10^5

Polyclonal ELSIA signals confirmed that specific binding of GFPrft displaying phage only occurred in the population received after the 2nd round of panning (Figure 4.19).

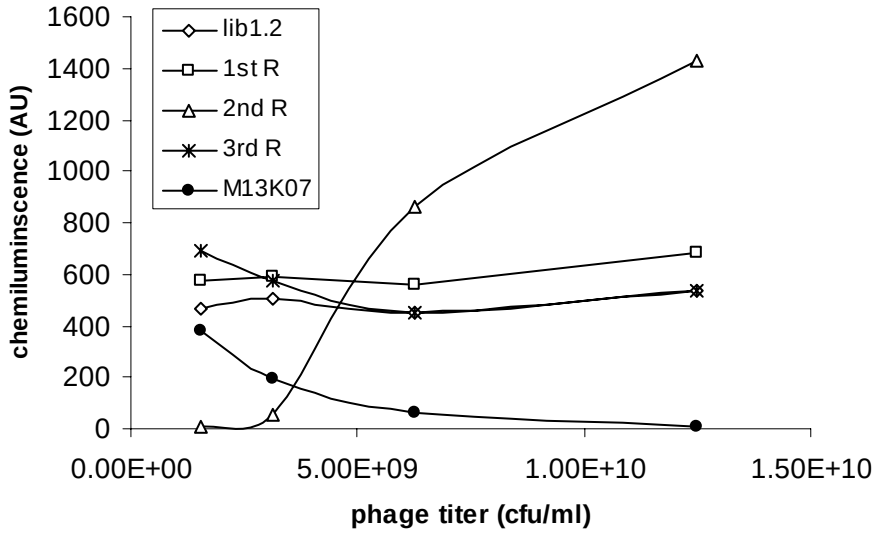


Figure 4.19. Polyclonal phage ELISA signals from phage recovered after the indicated panning round. Plates were coated with transferrin, phage was added as indicated and detected with anti M13 HRP antibody via chemiluminescence.

4.3.4.2 Single clone ELISA analyses

We tested around 200 colonies from the second round of panning. Binding to glucose oxidase and laccase was reduced in general. We identified several clones with

increased binding affinity for transferrin when compared to streptavidin (data not shown and Figure 4.20).

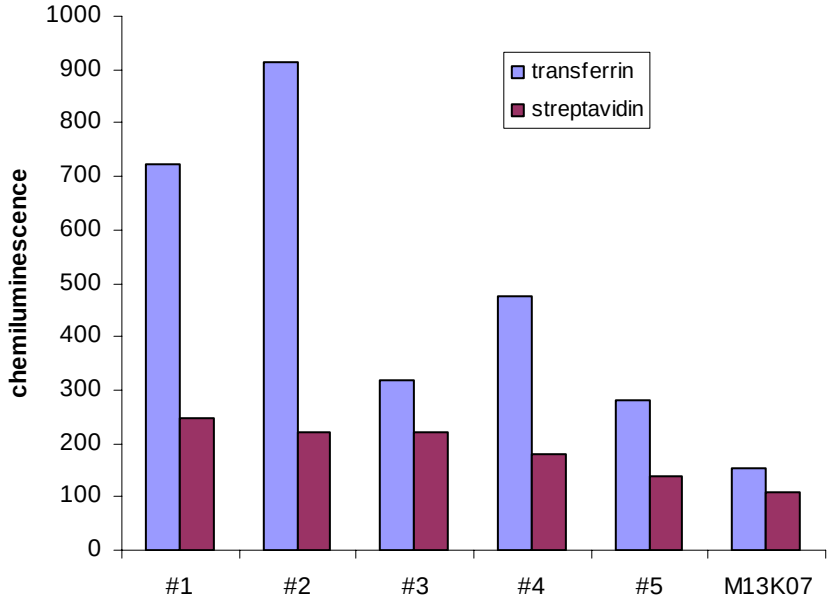


Figure 4.20. Positive single clone ELISA showing specificity towards transferrin. Clones were obtained from 2nd round of panning using bead approach

However, upon further amplification of the phage carrying the recovered phagemid and PEG precipitation, specific binding of the purified phage could not be confirmed by ELISA.

Since we did not identify any transferrin binder in our single clone ELISA screens, the question arose whether the phage concentration in the supernatant of the 96-well plates was not enough to detect low affinity binders to transferrin. To exclude the possibility that too low phage titers were achieved, we performed an ELISA in which we coated plates with serial dilutions of supernatant and compared it known phage titers. Phage binding was confirmed by the use of anti M13 antibody. ELISA signals indicated that at

least half of the 10 analyzed clones had titers around 10^8 cfu/ml whereas the other half contained around 10^7 cfu/ml (data not shown). Even with low titers such as 10^7 cfu/ml, binding to transferrin should be detected if the affinity had been at least in the micromolar range. Hence the lack of an ELISA signal is not due to low phage titers. It is plausible that not enough GFPrft was displayed on the phage after precipitation or the binding affinities or the binding affinity to transferrin is below 10^{-6} M.

4.4 CONCLUSION

We constructed a fluorescent GFP library containing a random sixmer amino acid insertion. The theoretical diversity of the sixmer peptide library is 6.4 x 10^7 individual sequences. Our library contained 10^7 GFP variants clones, 18% out of which were fluorescent. We were able to display fluorescent GFPrft containing a peptide insertion on filamentous phage particles using Tat mediated export. We further exposed the GFPrft library displayed on phage to several rounds of panning against ligands of interest and detected binding affinities of polyclonal phage populations. However, we could not isolate individual GFPrft variants that exhibited specific binding to lysozyme, streptavidin or transferrin.

Previously Paschke et al. reported the insertion of p24 peptide tag into GFP and Paramban et al. described the insertion of a 6 histidine residue sequence (Paramban et al. 2004; Paschke and Hohne 2005). However, both studies demonstrated that the GFP variants exhibited decreased fluorescence and low solubility. Our choice of the *E. coli* expression host DHF as well as the usage of the thermostable GFPrft variant as our scaffold allowed us to achieve higher cellular fluorescence. Further, this is the first

demonstration of the incorporation a peptide library into a functional GFP containing its original fold. We established the conditions which allowed the display of completely folded and fluorescent GFPrft on a filamentous particle with and without insertions.

Surprisingly, when we analyzed the fluorescence of phage particles displaying either GFPrft or GFPrft containing a sixmer insert, we found that all had a similar fluorescence, which is in contrast to the cellular fluorescence measurements by flow cytometry. In other words, the difference in the cellular fluorescence level did not seem to affect the incorporation of fluorescence GFP into the phage particle under the conditions used for phage assembly.

When phage particles displaying the GFPrft peptide library were subjected to two rounds of panning against a ligand of interest, we saw enrichment evident by an increase in the number of phage clones emerging sequential panning steps. We further demonstrated concentration-dependent binding of the recombinant polyclonal phage to target by ELISA analyses. However, further rounds of panning did not result in the enrichment of clones with higher binding affinities, even though more colony forming units were recovered. Despite the application of several washing steps, it appears that several unspecific phage particles were recovered after each round. If these unspecific variants exhibited lower cellular toxicity, their growth advantage over other variants would result in an over-representation of non-specific variants in the purified phage solution.

Altogether, we applied 3 different strategies to isolate GFP binders and to eliminate unspecific binders. After we were not able to enrich for binders using immunotubes and extensive washing, we included pre-absorption steps using immunotubes coated with an unrelated protein (glucose oxidase) combined with alterations of the blocking agent. Despites these efforts, we could not isolate any specific

binder. We therefore utilized a magnetic, streptavidin-coated bead display system for the immobilization of biotinylated transferrin combined with pre-absorption steps, changing the blocking agent used for each round of panning and increasing the stringency of the washing conditions for each round. Again, we still were not successful in isolating any ligand-binding GFP variant, even though polyclonal phage ELISA analyses showed low binding affinities. We therefore conclude that the binding affinities incorporated into GFPrft by a hexameric peptide insertion into the loop at position 157 did not produce high enough binding affinities to allow their separation from unspecific binding variants. Hence, the inclusion of another peptide sequence into a second loop of GFP could be desirable. Alternatively, the binding affinity of the individual phage particle could be strengthened by increasing the amount of displayed GFPrft. Our Tat-based phage display system was based on a 3+3 phage vector (see Introduction 1.7), which implies that not every pIII protein contains a Fos adaptor protein recruit to GFP molecule. The display of GFP is further assumed to be low since anchoring to the phage requires first the association of the adaptor proteins Jun and Fos, and the formation of disulfide bridges to introduce stabilizing covalent bonds between the adaptor molecules. At this time point, we are not able to determine the amount of GFP displayed per phage for which we would need further analyses. We also detected lower binding affinities of purified phage compared to culture supernatant. This could be either due to the effect of high salt used for phage precipitation which could interrupt the Jun/Fos adaptor system or the disruption of binding of the selected GFP variants to the target ligand. It is conceivable that disulfide bridge formation is not efficient enough to allow optimal GFP anchoring, even though the presence of disulfide bond was confirmed in both *E. coli* (data not shown, Tullman-Ercek, unpublished) and on phage particles (Crameri and Suter 1993; Paschke and Hohne

2005). It is possible that an optimization of the disulfide bridge formation step could increase the display of GFPrft particularly on precipitated phage.

In summary, we generated a fluorescent GFPrft library displaying 1.8 x 10^6 different peptide sequences. We were able to enrich polyclonal phage population that showed concentration-dependent binding to the ligand of choice. However, we believe that the binding affinities introduced into GFPrft are not high enough to allow the identification of single GFPrft variants. A second loop insertion and the change of phage vectors could help to achieve this goal.

Chapter 5

Conclusion and future directions

The work described here focused on both the biotechnological applications of the twin-arginine transporter pathway as well as the understanding of the basic principles of its function. In the chapter 2 and 4, we describe two applications based on properties of the Tat pathway. In chapter 2 we developed a novel two-hybrid system which takes advantage of the inherent proof-reading mechanism of the Tat pathway for the detection of interacting proteins. In chapter 4, we sought to develop a Tat-based phage display system. The Tat-based phage display system enables proteins requiring folding in the cytoplasmic environment to be exported after their folding maturation and to then assemble on to the phage. In the 3rd chapter and in the appendix, we listed work aiming to understand the underlying molecular basis of the Tat pathway. In chapter 3, we describe the identification of several amino acid substitutions in TatC which allow the export of a non-functional signal peptide containing two lysine residues instead of the hallmark twin-arginine residues that are found in nearly all signal peptides. Suppressor mutations in TatC were isolated indicating that this protein is responsible for the recognition of the RR sequence in Tat signal peptides. Finally, we described in the appendix two methods to detect the topology of TatA, using GFP fusions or the incorporation of a non-canonical amino acid containing a chemically reactive group.

The new bacterial two-hybrid system for the detection of interacting proteins (Chapter 2) capitalizes on the folding quality control mechanism of the Twin Arginine

Transporter (Tat) pathway. The Tat export pathway is responsible for the membrane translocation of folded proteins, including proteins consisting of more than one polypeptide, only one of which contains a signal peptide ("hitchhiker export"). One protein (bait) was expressed as a fusion to a Tat signal peptide whereas the second protein (prey) was fused to a protein reporter that can conferred a phenotype only after export into the bacterial periplasmic space. Since the prey-reporter fusion lacks a signal peptide, it can only be exported as a complex with the bait-signal peptide fusion which is capable of targeting the Tat translocon. Using maltose-binding protein as a reporter, clones expressing interacting proteins can be identified on maltose minimal media or on MacConkey plates. In addition, we introduced the use of the cysteine disulfide oxidase DsbA as a reporter for Tat mediated export. Export of a signal peptide-prey:bait-DsbA complex into the periplasm allowed complementation of *dsbA*$^{-}$ mutants and restored the formation of active alkaline phosphatase, which in turn was detected by a chromogenic assay.

The Tat two-hybrid system can be used as a new tool to identify protein-protein interaction on a genomic scale by including two libraries as bait and prey, or it can be used to identify protein interaction partners for a protein of interest. Further, the Tat pathway can be utilized as a filter for selecting protein variants having higher solubility and faster folding. Therefore it should be possible to use the Tat two-hybrid system as a genetic screen for the isolation of better folding and solubility of protein pairs while still maintaining their binding affinity for each other. This can be a powerful tool for *in vivo* co-evolution of interacting protein pairs. For example, beta lactamase confers resistance

to antibiotics containing a beta-lactam ring. An inhibitor protein of beta-lactamase, BLIP (for beta lactamase inhibiting protein), has been identified (Paradkar et al. 1994). However, BLIP is highly insoluble and expression in bacteria results in low yields. The Tat-based two-hybrid system could be used to optimize the solubility of BLIP while still maintaining its binding affinity for beta lactamase. Error prone PCR mutagenesis could be applied to isolate more soluble, but functional protein variants of BLIP. Since the interaction of beta-lactamase and BLIP has been well characterized, it serves as a paradigm for the modeling of protein-protein interactions. It could be interesting to see what amino acid bias filtering through the Tat-pathway would introduce, when optimize for expression and binding.

Similarly, several antibodies which are isolated from eukaryotic cell lines such as hybridoma cells are often poorly expressed. The solubility of antibody fragments such as a scFv or Fab could be improved by applying the Tat two-hybrid system following mutagenesis. This optimization step can also have the advantage that the expression of this antibody fragment could be optimized for expression in *E. coli* which could result in significant cost reduction for biotechnological large scale purifications.

In the forth chapter, we describe the use of a Tat-based phage display system in which we optimized the conditions for the display of GFP variants containing a sixmer random peptide insertion. We were able to produce 2×10^6 fluorescent GFP variants containing a peptide insertion and displayed them on phage particles. We optimized the expression of the GFPrft variant for phage incorporation and were able to detect binding of polyclonal phage particles to model proteins namely streptavidin, lysozyme and

transferrin. However, so far we were not able to identify single clones which would specifically bind to the protein of interest. Our analysis indicated that the binding affinities might be too low to identify individual clones in our single clone screens. Hence, the binding affinities should be increased by inserting an additional peptide library into a second loop of GFP. Alternatively, the phage vector could be changed so that the linking adaptor protein, Fos, which anchors Jun-GFP to the pIII protein of the phage particle would be displayed on every pIII protein. This would allow multiple GFP molecules to bind simultaneously, which could increase the affinity of the complete phage particle. Further it is possible that the insertion into the loop at position 157 caused steric hindrance which prevented specific binding, therefore insertion into a different loop could be beneficial. Also by including a certain bias into the peptide insertion, such as including the flexible amino acid glycine at the beginning and end of the peptide insertion, could further improve the diversity of displayable peptide sequences while still allowing a fluorescent fold of GFP.

In the third chapter, we provide genetic evidence for TatC being the protein component of the translocon that recognizes the twin-arginine motif within a Tat signal peptide. *In vitro* studies have suggested that the TatBC complex serves as the receptor for signal peptides targeted for export via the Twin-Arginine Translocation (Tat) pathway (Alami et al. 2003). Substitution of the hallmark twin-arginine dipeptide within a Tat signal peptide with two lysine residues abrogates export of physiological substrates in all organisms. We isolated and characterized several suppressor mutations that allow export

of a ssTor(KK)-GFP-SsrA tripartite fusion. We identified two amino acid suppressor mutations in the first cytoplasmic loop of TatC. In addition, two other amino acids in the first cytoplasmic loop exhibit epistatic suppression. Surprisingly, we also identified a suppressor mutation predicted to lie within the second periplasmic loop of TatC, a region that is not expected to interact directly with the signal peptide. The suppressor mutations allowed export of the native *E. coli* Tat substrate TorA with a KK substitution in its signal sequence. The cytoplasmic suppressor mutations conferred SDS sensitivity and partial filamentation indicating that Tat export of authentic substrates was impaired.

Having identified only amino acid substitutions within the N-terminal half of TatC and knowing that only the N-terminal part is interacting with the signal peptide as parallel *in vitro* cross-linking studies indicated, raises the question how much of the C-terminal half is essential for Tat functionality. Truncation analyses have been performed on TatA and TatB but not on TatC. The bacterium *Providencia stuartii* contains a TatC molecule that contains only the first half of TatC. The complete genome of *P. stuartii* has not been sequenced so far and a possible second homologue could be present. However, recent analysis of the Tat pathway in *P. stuartii* revealed that the deletion of this TatC variant abolishes export (Stevenson et al. 2007). The similarity of the N-terminal half of the *E. coli* with the TatC homologue of *P. stuartii* is 91%, and 75% of the amino acid residues are identical. Hence it is conceivable that only the N-terminal half is essential for functionality. To approach this question, we generated a truncation library of TatC containing 3.5×10^5 variants. When verifying the quality of this library, we identified variants of the sizes of 240 bp up to complete length, with the latter being represented by

8% of the library which we will analyze for functionality. Growth on SDS medium is a simple and fast way to test the functionality of the Tat pathway due to the Tat-dependent translocation of the amidases AmiA and AmiC. Therefore, it will be promising to investigate into the minimum length of TatC.

Finally, we describe in the appendix two methods to analyze the topology of TatA. Insights into the conformational dynamics of TatA could possibly shed light on the mechanism of the actual translocaton step. TatA is the protein component of the Twin-Arginine Transporter pathway that is most likely mediating the actual translocation step of a substrate through the membrane. Its topology however is still disputed. Two independent studies claim a dual topology for the C-terminus of TatA. However, they contradict each other in the location of the N-terminus which contains the only transmembrane domain of the protein. By monitoring GFP fluorescence of TatA-GFP fusions and by proteinase K analyses we can confirm the dual topology of the C-terminus providing evidence for both cytoplasmic and periplasmic localization. Further, we incorporated the non-canonical amino acid azidohomoalanine *in vivo* which allowed us to incorporate an affinity tag site-specifically into TatA. The newly introduced biotin-PEO tag can then be detected by a streptavidin-fluorophore conjugate that was applied from the periplasmic site of the cytoplasmic membrane. Using this approach, we confirmed that the C-terminal end of the amphipathic helix of TatA is accessible from the periplasmic site of the membrane. We demonstrated that the N-terminus of TatA is only accessible when Mg(II) is present during the reaction of the azido group of AHA with biotin-PEO-propargylamide

Recently, it was demonstrated that the TatA homologue of *Providencia stuartii* requires a posttranslational modification to be active. Its TatA homologue contains an 8

amino acid long N-terminal extension which inactivates the Tat pathway. Proteolytic cleavage by the quorum sensing-responsive rhomboid protease, AarA, eliminates the N-terminal 8 amino acid sequence rendering TatA active. It is remarkable that a short extension of TatA controls the functionality of the Tat pathway. There are two possible explanations for the inactivation of TatA which would be interesting to analyze. One plausible explanation could be that the addition of these extra 8 amino acids prevents the oligomerization of TatA with either itself or other protein components of the Tat pathway such as TatB or TatC. Another possibility could be that the extension, which introduces a strong hydrophilic character with a glutamate-serine-threonine sequence, could impose a conformational trap onto TatA. After the negative outside, positive inside rule, the transmembrane helix would now be absolutely forced into one topological orientation with the N-terminus facing the periplasmic site. This could either result in the inability of TatA to oligomerize or simply eliminate its flexibility within the membrane which might be required for protein translocation. We therefore think that the two methods described in the appendix could be useful to understand the dynamics of TatA by further analyzing an N-terminally extended TatA combined with co-immunopreciptation analyses.

Appendix

A.1 Topological dynamics of TatA, mechanistic insights into flipping membrane proteins

A.1.1 Summary

TatA is the protein component of the Twin-Arginine Transporter pathway that is presumed to mediate the actual translocation step through the membrane. Its topology however is still disputed. Two independent studies claim a dual topology for the C-terminus of TatA. However, they contradict each other regarding the location of the N-terminus which bears the only transmembrane domain. Here we describe studies to evaluate the dual topology of the C-terminus. We were further able to show that the N-terminus is accessible from the periplasm in the presence of Magnesium.

We developed two methods to monitor the topology of TatA: In the first approach we utilized a C-terminal GFP fusion. GFP is only fluorescent if it folds in the cytoplasm of the cell. We are able to show that the GFP is fluorescent as a TatA fusion, but on the other hand it is accessible from the periplasm, as proteinase K studies showed. This implies that GFP did fold in the cytoplasm but probably due to the flipping of the amphiphatic helix of TatA, localized at the periplasmic site of the cytoplasmic membrane. In a second approach, we utilized the site-specific incorporation of a non-canonical amino acid, the methionine surrogate azidohomoalanine (AHA). Using "click chemistry", AHA was covalently linked to biotin which could be detected by membrane impermeable streptavidin linked to a fluorescent dye. We were able to determine the location of the N-terminus and the end of the amphipathic helix. This is further the first incorporation of an artificial amino acid into cytoplasmic integral membrane protein *in*

vivo. We found that the export of Tat substrates increases the amount of TatA in the cytoplasmic membrane and that overexpression of Tat substrate encoded in an inducible plasmid saturates the translocon which represses the translocation of indigenously lower expressed Tat substrates.

This project has been performed in collaboration with Dr. A.J. Link.

A.1.2 Introduction

As summarized by Rapp and coworkers, the orientation of a transmembrane domain is simply determined by the charge amino acid residues adjacent to the transmembrane domain. Typically negatively charged residues face an exocytoplasmic compartment whereas positively charged residues are located in the cytosol (Rapp et al. 2007). TatA has one transmembrane domain which has only two positively charged residues at the C-terminal site of the α-helix. The N-terminal site is mainly hydrophobic with the exception of one glutamate residue. TatA can be found in high molecular weight oligomers in the cytoplasmic membrane. Electron microscopy studies revealed that TatA oligomers form asymmetric pore-like structures of varying sizes that are large enough to accommodate export of various Tat substrate (Gohlke et al. 2005). Recent NMR analysis on the TatA homologue from *Bacillus subtilis* revealed that the transmembrane domain is tilted at a 17˚ angle with respect to the membrane (Muller et al. 2007).

As mentioned above, the existing topology studies are contradictory. A total of 6 different TatA topologies are possible. Various studies have presented data supporting 4 of these topologies. Considering the physiological role of TatA, the fact that it is dynamically assembled and that it mediates the translocation of folded proteins of up to 180 kD in size, it is likely that its conformation requires certain flexibility. The possibility

that all 6 possible topologies are assumed at some point during the translocation cycle cannot be excluded (Figure A.1).

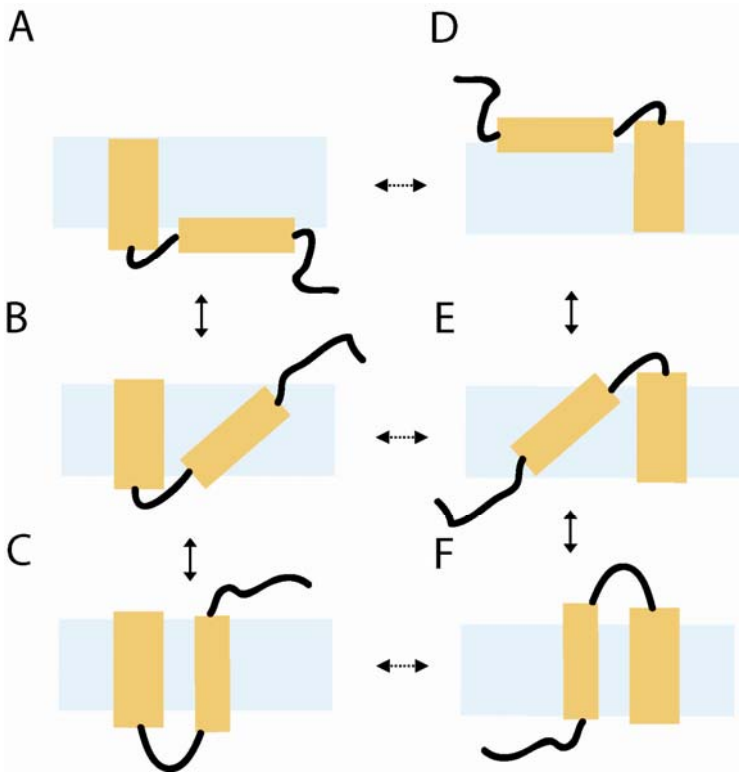


Figure A.1. Possible conformation of TatA

A.1.3 Results

A.1.3.1 TatA-GFP fusion analyzed by flow cytometry

Using C-terminal fusions to the reporters alkaline phosphatase and β-glucuronidase, Gouffi and coworker proposed that the N-terminal of TatA faces the periplasm whereas the C-terminus has a dual topology (Gouffi et al. 2004). Chan and coworkers however proposed a dual topology with the N-terminus facing the cytoplasm

based on analysis of the accessibility studies using membrane permeable and impermeable thio-reactive probes (Chan et al. 2007).

As was discussed in chapter 1, GFP expressed in *E. coli*, can only obtain its fluorescent fold in the cytoplasmic environment. When exported via the Sec pathway, GFP folds within the periplasm, but it is not able to form its chromophore and therefore it is non-fluorescent (Feilmeier et al. 2000). This property of GFP was exploited in a large scale study to determine the location of the C-termini of 601 transmembrane proteins (Daley et al. 2005). Membrane proteins are integrated into the cytoplasmic membrane via the SRP pathway (Luirink et al. 2005) or the Sec independent YidC pathway (Dalbey and Kuhn 2004). When fused to the C-terminus of a membrane protein, the fluorescent state of GFP is a direct indicator of the location of the C-terminal end of the protein: cell fluorescence can only be detected if GFP can fold in the cytoplasm indicating the cytoplasmic location of the fusion junction. We generated a fusion of GFP to the C-terminal of TatA (Figure A.2). The cells were found to be strongly fluorescent. Since GFP can only fold in the cytoplasm, the C-terminus must at least temporary be located in the cytoplasm so that GFP can assume its fluorescent state (Figure A.3, column 2).

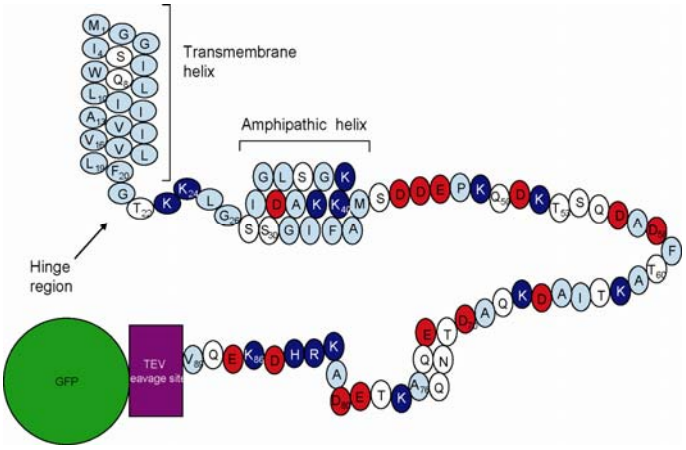


Figure A.2 C-terminal GFP fusion to TatA

Tat mutant strains are deficient in the export of 2 amidases (AmiA and AmiC), which cleave the peptide moiety of the peptidoglycan layer. It is the mislocalization of these two substrates which gives rise to the chain-forming phenotype of *tat* mutants. The chain formation can be quantitatively monitored by the forward scatter when using flow cytometry. When TatA is not expressed (A.3 column 1 versus 3) the population shifts in the forward/sideward scatter (FSC/SSC) profile due to chain formation. However, when the TatA-GFP fusion is coexpressed with large amounts of a Tat substrate (ssTorA-MBP), the FSC/SSC shows the same behavior as that for *tat* deficient strain indicating that the overexpressed Tat substrate out-competes the localization of presumably AmiA and AmiC (A.3 column 5). However moderate expression of a Tat substrate which solely relies on the expression from the leaky *trc* promoter, allow a different phenotype which can be seen as a shift in the FSC/SSC population which shifts towards w.t. cells population; at least based on the FSC/SSC profile, cells expressing a Tat substrate at lower levels resembled more the w.t. strain (Figure A.1.3 column 4). Further, we noticed an increase in fluorescence when co-expressing a Tat substrate which was directly proportional to the amount of the substrate expression (comparison Figure A.1.3 column 4 and 5). The increase of fluorescence could indicate that more TatA is stabilized in the cytoplasmic membrane when a Tat substrate is expressed.

We further wanted to know whether a high TatA expression level is dependent on the expression of other Tat proteins. We analyzed TatA-GFP expression together with the Tat substrate ssTorA-MBP in a *tatC* mutant strain. Surprisingly, when co-expressing TatA-GFP and ssTorA-MBP in a Δ*tatC* strain, fluorescence was still high (data not shown).

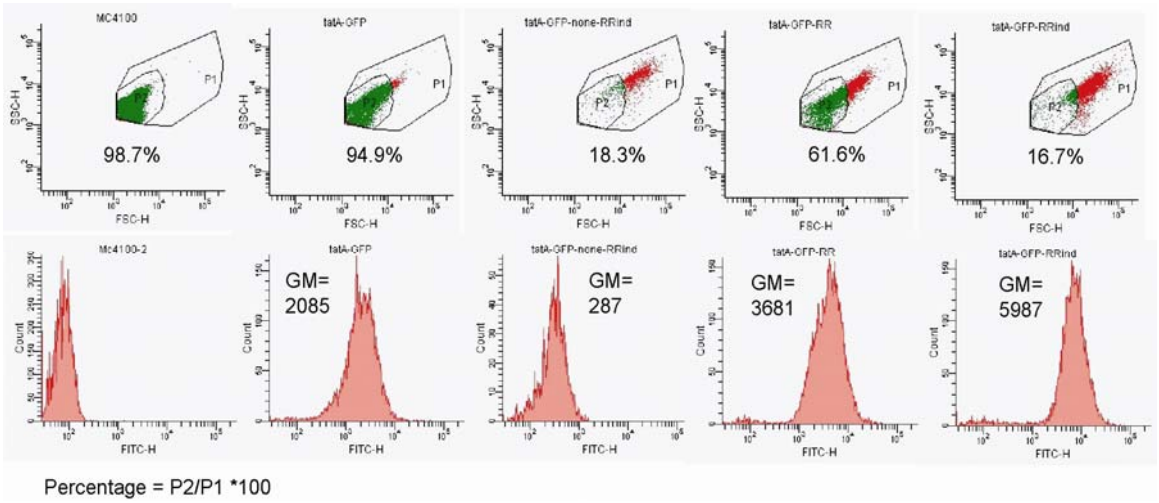


Figure A.3. GFP fluorescence of TatA-GFP with and without substrate co-expression. The first row shows the FSC/SSC profile whereas the second row shows GFP fluorescence at ex 488 nm/em 530 nm. Column 1: MC4100 cells, column 2 Δ*tatA/E* cells harboring the induced TatA-GPF fusion in pBAD33, column 3: same as column 2 but no induction, column 4: Δ*tatA/E* cells with induced TatA-GFP and uninduced ssTorA-MBP encoded in pTrc99 which has leaky expression, column 5: same as column 4, but TorA-MBP is induced with 0.1 mM IPTG.

To monitor the functionality of the TatA-GFP fusion in context of translocation of the co-expressed ssTorA-MBP fusion, we fractionated cells grown as listed above into periplasmic and cytoplasmic fractions and analyzed them by Western blot analysis using an anti-MBP antibody which was detected with a anti-mouse HRP conjugate (Figure A.1.4). The TatA-GFP fusion was able to rescue the Δ*tatA/E* deletion.

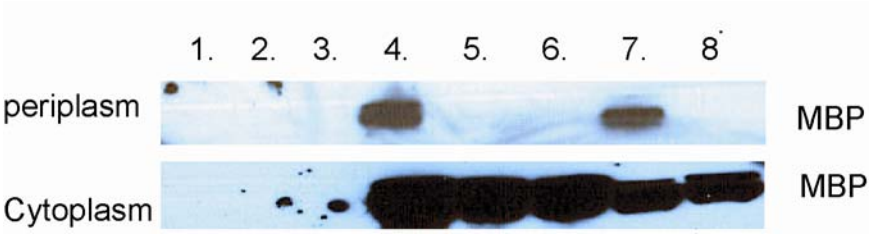


Figure A.4. Fractionation analysis of cells harboring a TatA-GFP fusion with and without the ssTorA-MBP substrate expressed. (1) MC4100 ,(2)TatA-GFP(induced), (3) TatA-GFP(ind) + pTrc99 empty, (4) TatA-GFP(ind.) + TorA-MBP(ind.), (5) TatA-GFP(not ind.) + TorA(RR)-MBP (ind.), (6) TatA-GFP(ind.) – TorA(KK)-MBP (ind.), (7) TatA-GFP(ind.) + TorA-MBP (no ind.), (8) TatA-GFP(ind.) + TorA(KK)MBP (no ind.)

A.1.3.2 Proteinase K studies – accessibility of the C-terminus from the periplasm

Since cells expressing TatA-GFP were fluorescent, we wanted to know whether GFP was also accessible from the periplasmic side. Cells cultivated as above, were spheroplasted by osmotic shock and lysozyme treatment resulting in the removal of the outer membrane and the peptidoglycan layer. We exposed the spheroplasted cells to 300 μg/ml of proteinase K for 30 min and analyzed them via Western blot analyses (Figure A.1.5). As a control for the intactness of the spheroplasts, we monitored the level of the cytoplasmic protein GroEL which was unaltered when treated with proteinase K. If the spheroplast were lysed, changes in the TatA-GFP proteins could also be due to proteinase K cleavage from the cytoplasmic side. As another control, we monitored the GroEL and GFP levels of cells that were lysed by treatment with 1% Triton-X. Our data demonstrates that GFP can be cleaved-off from the periplasmic side of the inner membrane indicating that the GFP does localize on that side of the membrane. When analyzing the accessibility of GFP in *ΔtatC* mutants, we further could confirm the cleavage of GFP. (Figure A.5). This implies that the periplasmic localization of the C-terminus of TatA is independent of its functionality.

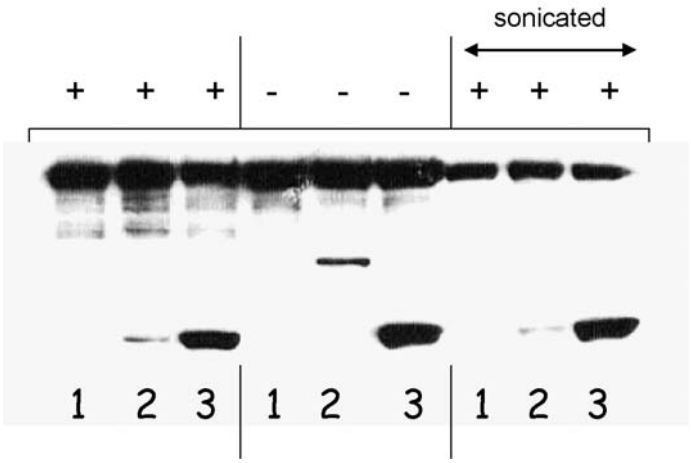


Figure A.5 proteinase K treatment of spheroplasts 3 h after induction. +/- indicates whether proteinase K was applied. (1) MC4100, (2) JARV16 (Δ*tatA*) harboring TatA-GFP, (3) MC4100 with GFP.

A.1.3.3 AHA incorporation into TatA, a chemical reaction introduces a new affinity tag

Originally, the topology of a membrane protein was assumed to be simple alpha helices crossing the membrane perpendicular for several times. Its conformation was simply defined based on which part enters and exits the membrane. 60 high-resolution membrane protein structures later, this view has changed. Membrane embedded helices can be long, short, kinked, half-inserted or crossing the membrane in different angles other than 90˚. Also, membrane protein complexes assemble more dynamically and spontaneously than assumed. The major disadvantage of currently used, protein-based topology markers is that they are bulky and their surface is charged. If fused to a membrane segment that localizes within the membrane or is kinked, the properties of the

topology marker might force the membrane domain to either site of the membrane, a bias which is determined by the properties of the marker protein.

We tried to circumvent this bias by incorporating the chemically reactive artificial amino acid azidohomoalanine (AHA). AHA is incorporated into proteins by exchanging methionine in the medium with AHA which causes the cell to charge their methionine tRNA with the surrogate. Hence, cells overexpressing the protein of interest in the artificial AHA medium incorporate AHA instead of a methionine in a site-specific manner. The reactive azido group can then be reacted in a Cu(I) catalyzed step with any propagylated ligand of interest. We utilized a triazole ligand together with a propargylated biotin containing a PEO spacer to introduce a new affinity tag (Figure A.6). The advantange of this approach is that biotin and its spacer are less bulky and presumeably less charged than any protein ligand. To identify the location of the incorporated biotin, we utilized fluorescently marked phycoerythrin (PE) streptavidin (488nm/575nm) applied to spheroplasted cells.

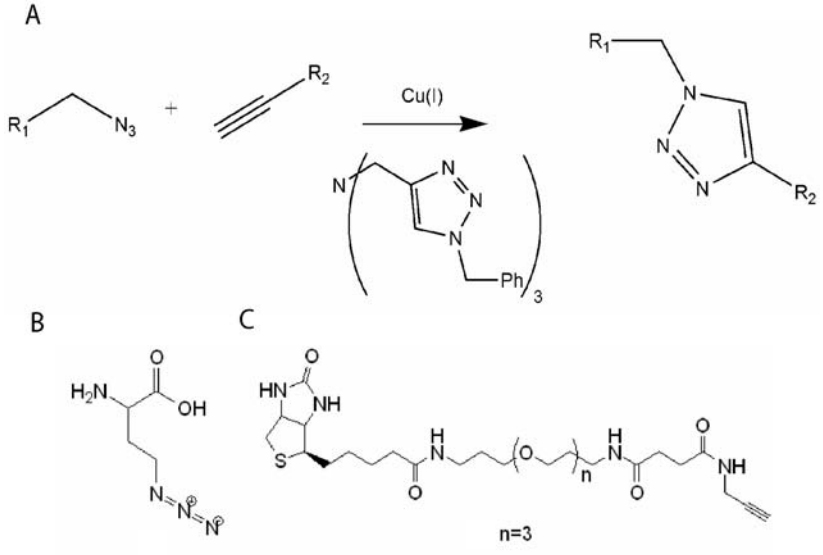


Figure A.6. Cu(I) catalyzed click chemistry. Reaction of an azide group with a propargyl group catalyzed by Cu(I) in the presence of a triazole ligand. (A) Overview of the click reaction, (B) as R1 we used the non-canonical amino acid azidohomoalanine, (C) biotin-PEO-propargylamide.

TatA contains a methionine at position 1 and 43. We eliminated the second methionine by site-directed mutagenesis giving rise to pTatA-M43L. Cells harboring a pBAD24 plasmid containing either TatA or the TatA(M43L) were grown in minimum media with and without Mg(II). After growth to the mid-logarithmic phase, the medium was switched to the inducing medium containing all amino acids with the exception that methionine was exchanged with AHA. Further, the glucose carbon source was exchanged with arabinose. After 4 h of growth in the induction media, cells were normalized and reacted at 4˚C with 40 μg/ml triazole ligand and an excess of biotin-PEO-propargylamide and Cu(I). After a 12 h reaction time, cells were spheroplasted using the same osmotic shock/lysozyme treatment as described and incubated with streptavidin-PE at room temperature for 1 h. Cells were then analyzed by flow cytometry to see whether streptavidin-PE did bind to the surface exposed biotin which was incorporated into the TatA variants (Figure A.7).

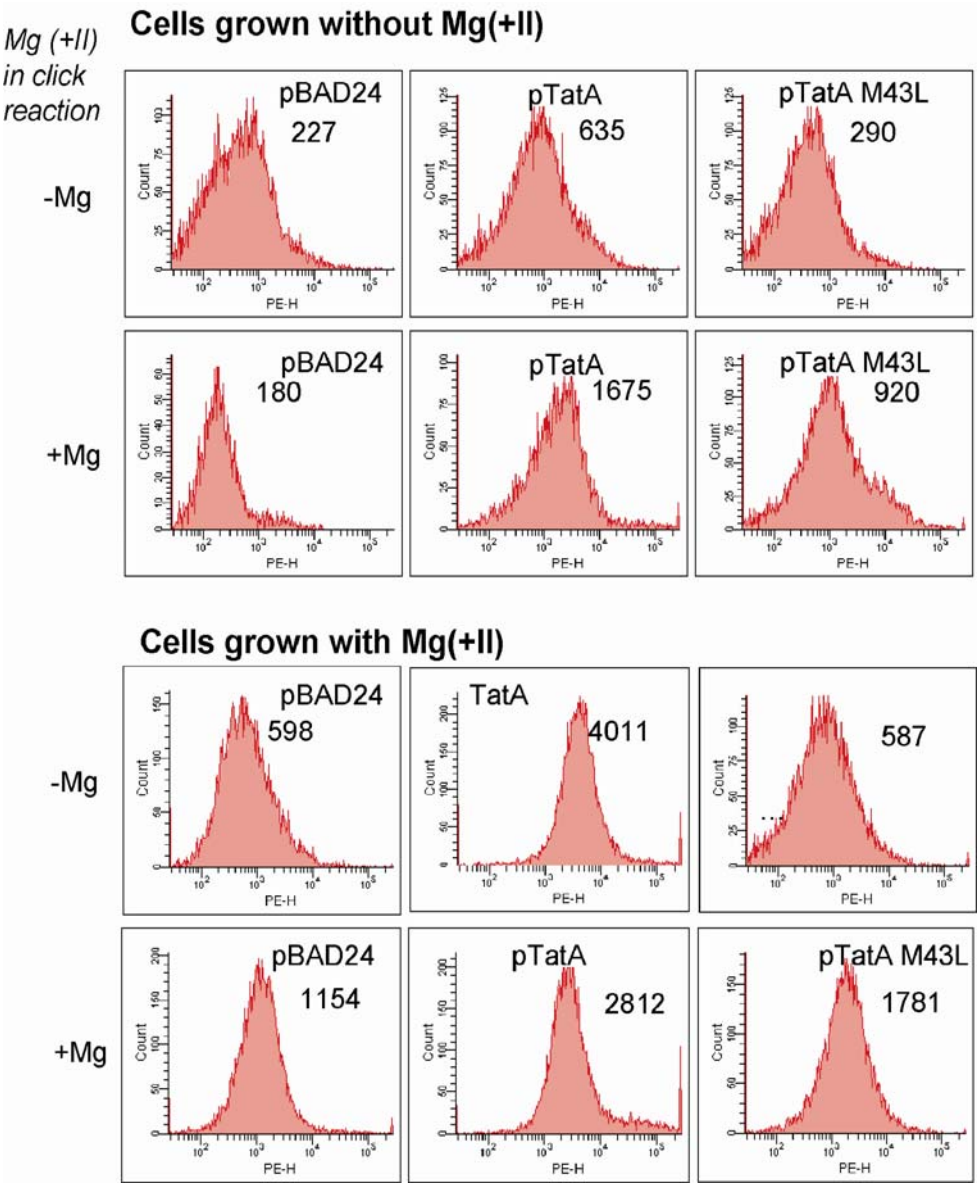


Figure A.7. Flow cytometry analysis of streptavidin-PE (488nm/575nm). Cells grown with azidohomoalanine(AHA) and reacted with biotin-PEO-propargylamide. After spaehroblastin, biotin was detected with streptavidin-PE.

To verify that the signal obtained was due to the expression of TatA and not any other protein such as beta lactamase which is the antibiotic resistance marker of the

plasmid and which localizes in the periplasm, we compared the TatA samples to cells containing induced empty plasmid. Our concerns were disproved, since only background noise was generated by cells containing the empty plasmid.

Our data indicates that the N-terminal methionine (or here AHA) was not accessible from the periplasmic side when there was no Mg(II) in the click reaction solution. However, whenever Mg(II) is depleted in the click reaction, the N-terminal AHA cannot be labeled/detected by the streptavidin-PE probe even though it is biotinylated (Figure A.8). To analyze the efficiency of the biotin incorporation and to ensure that differences seen where not due to variations in the protein expression levels of both TatA and TatA M43L, we performed a western blot analysis using streptavidin conjugated to HRP and anti-TatA antibodies detected with anti-rabbit HRP antibodies. The western blot result indicates that the Mg(II) addition does not interfere with the click reaction since we detected biotinylation (Figure A.8 and data not shown). Therefore it is possible that the effect of Mg(II) takes places on a cellular level, possibly affecting the ion gradient. Regardless of the N-terminal labeling of TatA we can confirm that the methionine at position 43 contributes a strong signal which indicates a periplasmic location of the end of the amphipathic helix.

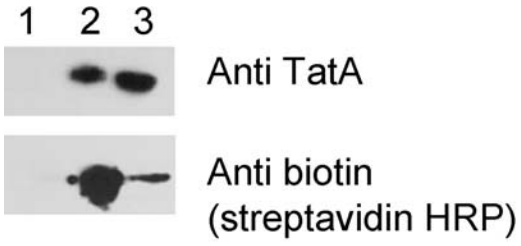


Figure A.8. Western blot analysis of cells grown without Mg(II) with 40 μg/ml AHA and 0.4% arabinose. Biotinylation was performed in the presence of Mg(II).

A.1.4 Conclusion and future perspective

In summary we can confirm that the C-terminus of TatA localizes in the periplasmic space, but we also demonstrated that the C-terminus is at least temporarily located in the cytoplasm since cells expressing the TatA-GFP fusions were fluorescent. It will be necessary to monitor how many of the GFP molecules, which were accessible to proteinase K, assumed a fluorescent fold. Together, this data confirmed the flexibility of the C-terminal end in respect to its localization. We can further point out that not only the C-terminal end of TatA localizes in the periplasm, we also detected that residue 43 of TatA is exposed to the periplasm. Residue M43, lies within the end part of the amphiphatic helix. By exchanging methionine with AHA, we introduced a biotin residue at this position which in turn we detected to be accessible from the periplasmic side as shown by labeling with a fluorescent streptavidin-PE probe. We further observed that the N-terminal methionine, or in our case AHA, at position 1 can be biotinylated, but it can be only detected with a streptavidin-PE probe when there was Mg(II) presence during the click reaction. Since the addition of Mg(II) to the reaction mixture did not affect the biotinylation, it is likely that Mg(II) affects the ion gradient across the cytoplasmic membrane which has consequences on the topology of TatA. We think it will be necessary to further test the functionality of TatA containing AHA to ensure that labeling and detection happen under physiological conditions. A pulse chase experiment with the overexpression of a Tat substrate will be useful to confirm the intactness of TatA containing AHA.

There are several advantages of using a chemical reactive artificial amino acid for topology studies. As already mentioned protein reporter fusions are bulky and charged, this can pull out the membrane region to either side of the membrane. The amino acid composition of periplasmic or cytoplasmic proteins diverges which might have an effect

onto which side the membrane domain is pulled. Our biotin-PEO-propargylamide is small and less surface-charged than any protein marker, which might circumvent this problem. Another advantage is AHA can be fused to an affinity tag, as well as being directly linked to fluorophores or a photo-reactive cross-linking agent. However, since currently the possible cross-linking reagent is not commercially available, it would have to be synthesized as described by (Hiramatsu et al. 2007).

Finally, we were able to confirm that the overexpession of a Tat substrate out-competes the translocation of indigenous Tat substrates. We noticed a shift of FSC/SSC towards the same values as a *tat* deletion strains when expressing ssTorA-MBP from the induced *trc* promoter. However this phenotype was less pronounced when less ssTorA-MBP was expressed. Also, the more substrate we expressed the higher the fluorescence, which indicates that more TatA is stabilized when the translocon is active. Surprisingly, we saw the same increase in fluorescence when overexpressing ssTorA-MBP in a Δ*tatC* strain. TatC is essential for the signal peptide recognition. Hence is it conceivable that TatA still senses the presence of a Tat substrate without the essential part of the signal recognition complex?

A.1.5 Material and methods

A.1.5.1 Plasmid construction and bacterial strains

Name	Genotype	Source
MC4100	*F⁻ lacΔU169 araD139 rpsL150 thi flbB5301 deoC7 ptsF25 relA1*	(Casadaban and Cohen 1979)
B1lk0	MC4100, Δ*tatC*	(Bogsch et al. 1998)
JARV16	MC4100, Δ*tatA*, Δ*tatE*	
M15MA	M15, Δ*metE*	(Link et al. 2003)
M15MAA	M15MA, Δ*tatA::kanFRT*	This study
pBAD33	pACYC origin, p_{BAD} promoter	(Guzman et al. 1995)
pTatA-GFP	pBAD33, *tatA-tev_cleavage site-gfp*	This study
pBAD24	pBR322 origin, p_{BAD} promoter	(Guzman et al. 1995)
pTatA	pBAD24 *tatA*	This study
pTatA-M43L	pBAD24 *tatA(M43L)*	This study

A.1.5.2 Growth conditions

A single colonies were grown overnight in M9 minimal medium supplemented with 40 μg/ml of all twenty natural amino acids and 100 μg/ml ampicillin. Cultures were subcultured 1/50 into fresh medium. Upon reaching an $O.D._{600}$ of 0.5, the cells were pelleted, washed twice with M9 salts and resuspended in M9 medium containing AHA instead of methionine and all remaining natural amino acids and 0.4% arabinose as a carbon source. After 4 h induction cells were harvested by centrifugation and normalized to $O.D._{600}$ of 0.5.

A.1.5.3 Flow cytometry

Cells were cultivated at 37˚C in LB after subculturing 1/50 from saturated overnight cultures. Cultures were induced in mid-logarithmic growth phase with 0.2%

arabinose, and after a 3 h induction period, cells were washed once with PBS, diluted 100 times and analyzed by flow cytometry on FACSAria. Successful sorting gates were set slightly above background (2 fold) based on fluorescence at 488 nm/530 nm.

For FACS analysis of the strepatvidin-PE labeling experiments, spheroplast were diluted 100 times into PBS and analyzed for fluorescence at 488 nm/575 nm.

A.1.5.4 Spheroplasting

Cultures were normalized to $O.D._{600}$ of 2, were washed once in 500 µl ice cold 50mM Tris pH 8. After centrifugation, cells were resuspend in 350 µl of 0.75M Sucrose-100mM Tris (pH 8). 700 µl of ice cold 1mM EDTA (pH 8) were added drop-wise while slowly vortexing the tube. 35 µl of lysozyme (20 mg/ml in 0.75M Sucrose-100 mM Tris, pH 8) were added and incubated for 10 min at room temperature. To stabilize the spheroplasts, 50 µl of ice cold 0.5 M $MgCl_2$ were added and incubate on ice for 10 min. After centrifugation, cells were resuspended in 1 ml PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 2 mM KH_2PO_4).

A.1.5.5 Proteinase K treatment

The 1 ml sphaeorblast solution was split into two aliquots of 490 µl each. One aliquot was incubated with 5 µl of Proteinase K (Proteinase K stock = 30mg/ml in 50mM Tris pH8) on ice, the second one was placed on ice as a negative control. After 1 h of incubation, 5 µl phenylmethanesulphonylfluoride (PMSF stock = 33mg/ml in isopropanol) were to both aliquots and samples were incubated for further 10 min on ice.

Typically 150 µl of treated spaehroblasts were combined with 50 µl SDS loading buffer and 22 µl were loaded onto 4%-20% Tris gels.

A.1.5.6 azidohomoalanine and Biotin-PEO-propargylamide synthesis and purification

Biotin-PEO-propargylamide was prepared by dissolving TFP-PEO-Biotin (Pierce, Rockford, IL) in excess neat propargylamine while stiring at room temperature. After 10 minutes, the solution was added dropwise to ethyl ether. A white precipitate formed and was collected by centrifugation. The precipitate was dried.

Synthesis and purification of AHA from N-α-Boc protected diaminopropionic acid was done as described previously (Link et al. 2007b; 2007a)

A.1.5.7 Click reaction and fluorescent labeling

Cells were normalize to a total $O.D._{600}$ of 2 and resuspended in 1 ml PBS (pH 7.4). About 40 µg triazole ligand was added together with approximately 50 µM of biotin-PEO-propargylamide and an excess amount of Cu(I). The concentration of biotin-PEO-propargylamide represented approximately a 100-fold excess of alkyne to azide. The reaction was allowed to proceed for 12-16 h at 4 ºC with agitation. At the conclusion of the reaction, the cells were pelleted and spheroplasted as described above.

200 µl of the spheroplast suspension were exposed to 1000 µg streptavidin-PE for 1 h while rotating at room temperature.

A.2 IDENTIFICATION OF NEW TAT SUBSTRATES BASED ON THEIR C-TERMINAL ANCHOR SEQUENCE

It seems striking that the five reported membrane-anchored Tat substrates are only inserted into the membrane by a short hydrophobic stretch at the C-terminus (Hatzixanthis et al. 2003). Most transmembrane proteins have either several transmembrane domains or they have a one-domain insertion at the N-terminus. Typically these proteins are facing the cytoplasmic side. Further, every protein marked for export contains a signal peptide at the N-terminus not at the C-terminus. Based on the phenomenon that only Tat substrates contain a C-terminal membrane anchor, the *E. coli* genome was searched for proteins which only contained a short C-terminal hydrophobic stretch that would be long enough to insert into the membrane. We utilized an algorithm which first determines the possible transmembrane domains of all proteins encoded in *E. coli* using the online transmembrane Hidden Markov model (TMHMM) program; a second program was written using Perl, which extracts all proteins containing only one transmembrane domain at the C-terminus. The number of proteins identified is surprisingly low. As a proof of principle, several Tat substrates out of the originally described enzymes were identified. We only report proteins longer than 100 amino acids (summary in Table A1). Possible functions were assigned based on BLAST and Smith-Waterman searches.

Table A1. Proteins containing a C-terminal hydrophobic stretch identified with search algorithm

Protein name	Amino acid length	End of transmembrane domain	(Possible) function
ybjT	486	470	putative oxidoreductase[a]
flk	331	330	flagellar assembly protein
ygaM	113	110	putative inner membrane protein[a]
ygdL	268	257	putative enzyme, NAD(P)-binding[a]
hybO[b]	372	354	
ygiM	206	195	arylsulfatase inVibrio fischeri ES114[a]
rfaJ	338	332	UDP-D-glucose:(galactosyl)lipopolysaccharide glucosyltransferase
nrfF[c]	127	121	part of formate-dependent nitrite reductase[a]
fdnH[b]	294	279	
ybeS	475	473	putative enzyme of polynucleotide modification; putative chaperone (J-protein like)

b: has been already identified as a Tat substrate

c: Tat substrate not further analyzed in respect of the export

a: seemed to be interesting functions

The proteins ybjT, ygdL, rfaJ, ybeS were of particular interest since they do not contain signal peptides and seem to either associated with a membrane or redox cofactors. The fact that they were missing a signal peptide was especially interesting, since if they are Tat substrates, they would require to use the hitchhiker mechanism through the Tat pore.

Proteinase K studies, performed as described under section A.1, confirmed that YgdL and YbjT face the periplasmic side of the membrane whereas YgaM is not accessible from the periplasmic side. Further analysis would be necessary to determine their Tat dependency.

A.3 A NEGATIVE GENETIC SCREEN FOR THE IDENTIFICATION OF FACTORS INVOLVED IN TAT EXPORT

Objective: identification of proteins that promote Tat export.

Introduction: The chloramphenicol resistance is transferred to a bacterium by the gene encoding chlorampenicol acetyl transferase (CAT). CAT transfers an acetyl-CoA unit onto chloramphenicol which results in resistance to this antibiotic. Since acetyl-CoA is only available in the cytoplasm, cat is only functional in this subcellular localization. Hence, fusion of the cat gene to a Tat signal peptide provides a genetic screen for a loss-of-function (Hicks et al. 2005).

Results: We generated a transposon library in MC4100 ($2 \times 10^{5)}$ which we screened with a ssTat-Cat reporter fusion for growth on chloramphenicol plate. If this construct is exported, the cell cannot grow. To avoid the isolation of insertions into any of the *tat* genes, we expressed the complete *tat* operon from a second plasmid. Transposon insertion into MurD, a UDP-N-acetylmuramoylalanine-D-glutamate ligase, which catalyzes the addition of the second amino acid to the peptide moiety of the monomer unit of peptidoglycan was pulled out several times. However, this is most likely not related to the Tat pathway since this enzyme localizes on the periplasmic side of the membrane and would hardly be able to promote Tat mediated export. It seems plausible that deletion would confer a higher resistance to chloramphenicol by either changing the properties of the peptidoglycan in a way that less chloramphenicol could enter the cell or by including chloramphenicol into the peptidoglcyan biogenesis pathway to compensate the intermediates that accumulate due to the MurD deletion.

Authors's list of publications

Strauch, E.M., and Georgiou, G. 2007a. A bacterial two-hybrid system based on the twin-arginine transporter pathway of E. coli. *Protein Sci* **16:** 1001-1008.

Strauch, E.M., and Georgiou, G. 2007b. Escherichia coli tatC Mutations that Suppress Defective Twin-Arginine Transporter Signal Peptides. *J Mol Biol*. **374: 283-291**

Glossary

aa amino acid

ATP adenosine triphosphate

BLAST basic local alignment search tool

BSA bovine serum albumin

cfu colony forming units

cpTat chloroplast twin-arginine translocation pathway

DMSO dimethylsulfoxide

DNA deoxyribonucleic acid

ELISA enzyme-linked immunosorbent assay

F_{AB} antibody fragment

FACS fluorescence-activated cell sorting

FLAG tag peptide sequence D-Y-K-D-D-D-D-K

GFP green fluorescent protein

HRP horseradish peroxidase

IgG immunoglobulin

IPTG isopropyl-β-D-thiogalactopyranoside

kDa kilo Dalton

KK two lysine residues

LB Luria-Bertani (Miller) growth medium

MBP maltose binding protein

MOPS 3-(N-morpholino)propane sulfonic acid

NADPH Nicotinamide adenine dinucleotide phosphate hydrogen

OD optical density

PAGE Polyacrylamide gel electrophoresis

pNPP para-nitrophenyl phosphat

PBS/PBST Phosphate buffered saline/PBS tween

PCR Polymerase chain reaction

PhoA alkaline phosphatase

PEG polyethylene glycol

PMF proton motive force

RR two arginine residues

SDS Sodium dodecyl sulfate

scFv single-chain variable fragment

Sec The general secretory pathway

SRP Signal recognition particle

ss signal sequence/signal peptide

Tat Twin arginine translocon

TBS/TBST tris-buffered saline, tris-buffered saline with Tween-20

TMAO trimethylamino-N-oxide

TorA trimethylamino N-oxide reducatase

$V_{H/L}$ variable domain of the heavy (H) or light (L) chain of an antibody

w.t. wild type

References

Abate, C., Luk, D., Gentz, R., Rauscher, F.J., 3rd, and Curran, T. 1990. Expression and purification of the leucine zipper and DNA-binding domains of Fos and Jun: both Fos and Jun contact DNA directly. *Proc Natl Acad Sci U S A* **87:** 1032-1036.

Akiyama, Y., and Ito, K. 1993. Folding and assembly of bacterial alkaline phosphatase in vitro and in vivo. *J Biol Chem* **268:** 8146-8150.

Alami, M., Luke, I., Deitermann, S., Eisner, G., Koch, H.G., Brunner, J., and Muller, M. 2003. Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli. *Mol Cell* **12:** 937-946.

Alder, N.N., and Theg, S.M. 2003. Energetics of protein transport across biological membranes. a study of the thylakoid DeltapH-dependent/cpTat pathway. *Cell* **112:** 231-242.

Allen, S.C., Barrett, C.M., Ray, N., and Robinson, C. 2002. Essential cytoplasmic domains in the Escherichia coli TatC protein. *J Biol Chem* **277:** 10362-10366.

Austermuhle, M.I., Hall, J.A., Klug, C.S., and Davidson, A.L. 2004. Maltose-binding protein is open in the catalytic transition state for ATP hydrolysis during maltose transport. *J Biol Chem* **279:** 28243-28250.

Bageshwar, U.K., and Musser, S.M. 2007. Two electrical potential dependent steps are required for transport by the Escherichia coli Tat machinery. *The Journal of cell biology* **179:** 87-99.

Barrett, C.M., Freudl, R., and Robinson, C. 2007. Tat-dependent export in the apparent absence of tatABC or TatA complexes using modified Escherichia coli TatA subunits that substitute for TatB. *J Biol Chem*.

Behrendt, J., Standar, K., Lindenstrauss, U., and Bruser, T. 2004. Topological studies on the twin-arginine translocase component TatC. *FEMS Microbiol Lett* **234:** 303-308.

Berkmen, M., Boyd, D., and Beckwith, J. 2005. The nonconsecutive disulfide bond of Escherichia coli phytase (AppA) renders it dependent on the protein-disulfide isomerase, DsbC. *J Biol Chem* **280:** 11387-11394.

Berks, B.C. 1996. A common export pathway for proteins binding complex redox cofactors? *Mol Microbiol* **22:** 393-404.

Berks, B.C., Palmer, T., and Sargent, F. 2003. The Tat protein translocation pathway and its role in microbial physiology. *Adv Microb Physiol* **47:** 187-254.

Bernhardt, T.G., and de Boer, P.A. 2003. The Escherichia coli amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. *Mol Microbiol* **48:** 1171-1182.

Bessette, P.H., Aslund, F., Beckwith, J., and Georgiou, G. 1999. Efficient folding of proteins with multiple disulfide bonds in the Escherichia coli cytoplasm. *Proc Natl Acad Sci U S A* **96:** 13703-13708.

Bjedov, I., Tenaillon, O., Gerard, B., Souza, V., Denamur, E., Radman, M., Taddei, F., and Matic, I. 2003. Stress-induced mutagenesis in bacteria. *Science* **300:** 1404-1409.

Blaudeck, N., Kreutzenbeck, P., Freudl, R., and Sprenger, G.A. 2003. Genetic analysis of pathway specificity during posttranslational protein translocation across the Escherichia coli plasma membrane. *J Bacteriol* **185:** 2811-2819.

Blaudeck, N., Kreutzenbeck, P., Muller, M., Sprenger, G.A., and Freudl, R. 2005. Isolation and characterization of bifunctional Escherichia coli TatA mutant proteins that allow efficient tat-dependent protein translocation in the absence of TatB. *J Biol Chem* **280:** 3426-3432.

Bogsch, E.G., Sargent, F., Stanley, N.R., Berks, B.C., Robinson, C., and Palmer, T. 1998. An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem* **273:** 18003-18006.

Bolhuis, A., Mathers, J.E., Thomas, J.D., Barrett, C.M., and Robinson, C. 2001. TatB and TatC form a functional and structural unit of the twin-arginine translocase from Escherichia coli. *J Biol Chem* **276:** 20213-20219.

Boos, W., and Shuman, H. 1998. Maltose/maltodextrin system of Escherichia coli: transport, metabolism, and regulation. *Microbiol Mol Biol Rev* **62:** 204-229.

Bork, P., Jensen, L.J., von Mering, C., Ramani, A.K., Lee, I., and Marcotte, E.M. 2004. Protein interaction networks from yeast to human. *Curr Opin Struct Biol* **14:** 292-299.

Bowden, G.A., Baneyx, F., and Georgiou, G. 1992. Abnormal fractionation of beta-lactamase in Escherichia coli: evidence for an interaction with the inner membrane in the absence of a leader peptide. *J Bacteriol* **174:** 3407-3410.

Bradshaw, R.A., Cancedda, F., Ericsson, L.H., Neumann, P.A., Piccoli, S.P., Schlesinger, M.J., Shriefer, K., and Walsh, K.A. 1981. Amino acid sequence of Escherichia coli alkaline phosphatase. *Proc Natl Acad Sci U S A* **78:** 3473-3477.

Bronstein, P.A., Marrichi, M., Cartinhour, S., Schneider, D.J., and DeLisa, M.P. 2005. Identification of a twin-arginine translocation system in Pseudomonas syringae pv. tomato DC3000 and its contribution to pathogenicity and fitness. *J Bacteriol* **187:** 8450-8461.

Buchanan, G., Leeuw, E., Stanley, N.R., Wexler, M., Berks, B.C., Sargent, F., and Palmer, T. 2002. Functional complexity of the twin-arginine translocase TatC component revealed by site-directed mutagenesis. *Mol Microbiol* **43:** 1457-1470.

Buchanan, G., Sargent, F., Berks, B.C., and Palmer, T. 2001. A genetic screen for suppressors of Escherichia coli Tat signal peptide mutations establishes a critical role for the second arginine within the twin-arginine motif. *Arch Microbiol* **177:** 107-112.

Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., et al. 2005. Interaction network containing conserved and essential protein complexes in Escherichia coli. *Nature* **433:** 531-537.

Casadaban, M.J., and Cohen, S.N. 1979. Lactose genes fused to exogenous promoters in one step using a Mu-lac bacteriophage: in vivo probe for transcriptional control sequences. *Proc Natl Acad Sci U S A* **76:** 4530-4533.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. 1994. Green fluorescent protein as a marker for gene expression. *Science* **263:** 802-805.

Chan, C.S., Howell, J.M., Workentine, M.L., and Turner, R.J. 2006. Twin-arginine translocase may have a role in the chaperone function of NarJ from Escherichia coli. *Biochemical and biophysical research communications* **343:** 244-251.

Chan, C.S., Zlomislic, M.R., Tieleman, D.P., and Turner, R.J. 2007. The TatA subunit of Escherichia coli twin-arginine translocase has an N-in topology. *Biochemistry* **46:** 7396-7404.

Chang, C.N., Kuang, W.J., and Chen, E.Y. 1986. Nucleotide sequence of the alkaline phosphatase gene of Escherichia coli. *Gene* **44:** 121-125.

Clarke, A.J. 1993. Extent of peptidoglycan O acetylation in the tribe Proteeae. *J Bacteriol* **175:** 4550-4553.

Cline, K., Ettinger, W.F., and Theg, S.M. 1992. Protein-specific energy requirements for protein transport across or into thylakoid membranes. Two lumenal proteins are transported in the absence of ATP. *J Biol Chem* **267:** 2688-2696.

Cline, K., and Mori, H. 2001. Thylakoid DeltapH-dependent precursor proteins bind to a cpTatC-Hcf106 complex before Tha4-dependent transport. *The Journal of cell biology* **154:** 719-729.

Cody, C.W., Prasher, D.C., Westler, W.M., Prendergast, F.G., and Ward, W.W. 1993. Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. *Biochemistry* **32:** 1212-1218.

Cohen, G.N., and Rickenberg, H.V. 1956. Not Available. *Annales de l'Institut Pasteur* **91:** 693-720.

Cole, N.B., Smith, C.L., Sciaky, N., Terasaki, M., Edidin, M., and Lippincott-Schwartz, J. 1996. Diffusional mobility of Golgi proteins in membranes of living cells. *Science* **273:** 797-801.

Collet, J.F., and Bardwell, J.C. 2002. Oxidative protein folding in bacteria. *Mol Microbiol* **44:** 1-8.

Cormack, B.P., Valdivia, R.H., and Falkow, S. 1996. FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173:** 33-38.

Crameri, A., Whitehorn, E.A., Tate, E., and Stemmer, W.P. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat Biotechnol* **14:** 315-319.

Crameri, R., and Suter, M. 1993. Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for selection of functional gene products linked to the genetic information responsible for their production. *Gene* **137:** 69-75.

Cubitt, A.B., Heim, R., Adams, S.R., Boyd, A.E., Gross, L.A., and Tsien, R.Y. 1995. Understanding, improving and using green fluorescent proteins. *Trends Biochem Sci* **20:** 448-455.

Dabney-Smith, C., Mori, H., and Cline, K. 2006. Oligomers of Tha4 organize at the thylakoid Tat translocase during protein transport. *J Biol Chem* **281:** 5476-5483.

Dalbey, R.E., and Kuhn, A. 2004. YidC family members are involved in the membrane insertion, lateral integration, folding, and assembly of membrane proteins. *The Journal of cell biology* **166:** 769-774.

Daley, D.O., Rapp, M., Granseth, E., Melen, K., Drew, D., and von Heijne, G. 2005. Global topology analysis of the Escherichia coli inner membrane proteome. *Science* **308:** 1321-1323.

Daugherty, P.S. 2007. Protein engineering with bacterial display. *Curr Opin Struct Biol* **17:** 474-480.

de Leeuw, E., Granjon, T., Porcelli, I., Alami, M., Carr, S.B., Muller, M., Sargent, F., Palmer, T., and Berks, B.C. 2002. Oligomeric properties and signal peptide binding by Escherichia coli Tat protein transport complexes. *J Mol Biol* **322:** 1135-1146.

De Leeuw, E., Porcelli, I., Sargent, F., Palmer, T., and Berks, B.C. 2001. Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway. *FEBS Lett* **506:** 143-148.

de Lichtenberg, U., Jensen, L.J., Brunak, S., and Bork, P. 2005. Dynamic complex formation during the yeast cell cycle. *Science* **307:** 724-727.

de Marco, A. 2007. Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in Escherichia coli. *Nature protocols* **2:** 2632-2639.

DeLisa, M.P., Samuelson, P., Palmer, T., and Georgiou, G. 2002. Genetic analysis of the twin arginine translocator secretion pathway in bacteria. *J Biol Chem* **277:** 29825-29831.

DeLisa, M.P., Tullman, D., and Georgiou, G. 2003. Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway. *Proc Natl Acad Sci U S A* **100:** 6115-6120.

Denning, D.P., Patel, S.S., Uversky, V., Fink, A.L., and Rexach, M. 2003. Disorder in the nuclear pore complex: the FG repeat regions of nucleoporins are natively unfolded. *Proc Natl Acad Sci U S A* **100:** 2450-2455.

Derman, A.I., and Beckwith, J. 1991. Escherichia coli alkaline phosphatase fails to acquire disulfide bonds when retained in the cytoplasm. *J Bacteriol* **173:** 7719-7722.

Desvaux, M., and Hebraud, M. 2006. The protein secretion systems in Listeria: inside out bacterial virulence. *FEMS microbiology reviews* **30:** 774-805.

Drew, D., Sjostrand, D., Nilsson, J., Urbig, T., Chin, C.N., de Gier, J.W., and von Heijne, G. 2002. Rapid topology mapping of Escherichia coli inner-membrane proteins by prediction and PhoA/GFP fusion analysis. *Proc Natl Acad Sci U S A* **99:** 2690-2695.

Drew, D., Slotboom, D.J., Friso, G., Reda, T., Genevaux, P., Rapp, M., Meindl-Beinker, N.M., Lambert, W., Lerch, M., Daley, D.O., et al. 2005. A scalable, GFP-based pipeline for membrane protein overexpression screening and purification. *Protein Sci* **14:** 2011-2017.

Driessen, A.J. 1992. Precursor protein translocation by the Escherichia coli translocase is directed by the protonmotive force. *Embo J* **11:** 847-853.

Dubini, A., and Sargent, F. 2003. Assembly of Tat-dependent [NiFe] hydrogenases: identification of precursor-binding accessory proteins. *FEBS Lett* **549:** 141-146.

Economou, A., and Wickner, W. 1994. SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell* **78:** 835-843.

Facey, S.J., and Kuhn, A. 2004. Membrane integration of E. coli model membrane proteins. *Biochim Biophys Acta* **1694:** 55-66.

Facey, S.J., Neugebauer, S.A., Krauss, S., and Kuhn, A. 2007. The mechanosensitive channel protein MscL is targeted by the SRP to the novel YidC membrane insertion pathway of Escherichia coli. *J Mol Biol* **365:** 995-1004.

Feilmeier, B.J., Iseminger, G., Schroeder, D., Webber, H., and Phillips, G.J. 2000. Green fluorescent protein functions as a reporter for protein localization in Escherichia coli. *J Bacteriol* **182:** 4068-4076.

Fekkes, P., van der Does, C., and Driessen, A.J. 1997. The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *Embo J* **16:** 6105-6113.

Fields, S. 2005. High-throughput two-hybrid analysis. The promise and the peril. *Febs J* **272:** 5391-5399.

Fisher, A.C., and DeLisa, M.P. 2004. A little help from my friends: quality control of presecretory proteins in bacteria. *J Bacteriol* **186:** 7467-7473.

Fisher, A.C., Kim, W., and DeLisa, M.P. 2006. Genetic selection for protein solubility enabled by the folding quality control feature of the twin-arginine translocation pathway. *Protein Sci* **15:** 449-458.

Friden, P.M., Walus, L.R., Musso, G.F., Taylor, M.A., Malfroy, B., and Starzyk, R.M. 1991. Anti-transferrin receptor antibody and antibody-drug conjugates cross the blood-brain barrier. *Proc Natl Acad Sci U S A* **88:** 4771-4775.

Fromant, M., Blanquet, S., and Plateau, P. 1995. Direct random mutagenesis of gene-sized DNA fragments using polymerase chain reaction. *Anal Biochem* **224:** 347-353.

Galperin, M.Y., and Koonin, E.V. 2000. Who's your neighbor? New computational approaches for functional genomics. *Nat Biotechnol* **18:** 609-613.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., et al. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415:** 141-147.

Georgiou, G., and Segatori, L. 2005. Preparative expression of secreted proteins in bacteria: status report and future prospects. *Current opinion in biotechnology* **16:** 538-545.

Gerard, F., and Cline, K. 2006. Efficient twin arginine translocation (Tat) pathway transport of a precursor protein covalently anchored to its initial cpTatC binding site. *J Biol Chem* **281:** 6130-6135.

Gerard, F., and Cline, K. 2007. The Thylakoid Proton Gradient Promotes an Advanced Stage of Signal Peptide Binding Deep within the Tat Pathway Receptor Complex. *J Biol Chem* **282:** 5263-5272.

Gohlke, U., Pullan, L., McDevitt, C.A., Porcelli, I., de Leeuw, E., Palmer, T., Saibil, H.R., and Berks, B.C. 2005. The TatA component of the twin-arginine protein

transport system forms channel complexes of variable diameter. *Proc Natl Acad Sci U S A* **102:** 10482-10486.
Gon, S., Giudici-Orticoni, M.T., Mejean, V., and Iobbi-Nivol, C. 2001. Electron transfer and binding of the c-type cytochrome TorC to the trimethylamine N-oxide reductase in Escherichia coli. *J Biol Chem* **276:** 11545-11551.
Gouffi, K., Gerard, F., Santini, C.L., and Wu, L.F. 2004. Dual topology of the Escherichia coli TatA protein. *J Biol Chem* **279:** 11608-11615.
Graubner, W., Schierhorn, A., and Bruser, T. 2007. DnaK plays a pivotal role in Tat targeting of CueO and functions beside SlyD as a general Tat signal binding chaperone. *J Biol Chem*.
Gray, M.W., Lang, B.F., and Burger, G. 2004. Mitochondria of protists. *Annual review of genetics* **38:** 477-524.
Greene, N.P., Porcelli, I., Buchanan, G., Hicks, M.G., Schermann, S.M., Palmer, T., and Berks, B.C. 2007. Cysteine scanning mutagenesis and disulfide mapping studies of the TatA component of the bacterial twin arginine translocase. *J Biol Chem* **282:** 23937-23945.
Greenwood, J., Hunter, G.J., and Perham, R.N. 1991. Regulation of filamentous bacteriophage length by modification of electrostatic interactions between coat protein and DNA. *J Mol Biol* **217:** 223-227.
Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *J Bacteriol* **177:** 4121-4130.
Halbig, D., Wiegert, T., Blaudeck, N., Freudl, R., and Sprenger, G.A. 1999. The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of Zymomonas mobilis depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding. *Eur J Biochem* **263:** 543-551.
Harvey, B.R., Georgiou, G., Hayhurst, A., Jeong, K.J., Iverson, B.L., and Rogers, G.K. 2004. Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries. *Proc Natl Acad Sci U S A* **101:** 9193-9198.
Hatzixanthis, K., Clarke, T.A., Oubrie, A., Richardson, D.J., Turner, R.J., and Sargent, F. 2005. Signal peptide-chaperone interactions on the twin-arginine protein transport pathway. *Proc Natl Acad Sci U S A* **102:** 8460-8465.
Hatzixanthis, K., Palmer, T., and Sargent, F. 2003. A subset of bacterial inner membrane proteins integrated by the twin-arginine translocase. *Mol Microbiol* **49:** 1377-1390.
Heidrich, C., Templin, M.F., Ursinus, A., Merdanovic, M., Berger, J., Schwarz, H., de Pedro, M.A., and Holtje, J.V. 2001. Involvement of N-acetylmuramyl-L-alanine amidases in cell separation and antibiotic-induced autolysis of Escherichia coli. *Mol Microbiol* **41:** 167-178.
Hicks, M.G., Lee, P.A., Georgiou, G., Berks, B.C., and Palmer, T. 2005. Positive selection for loss-of-function tat mutations identifies critical residues required for TatA activity. *J Bacteriol* **187:** 2920-2925.

Hiniker, A., and Bardwell, J.C. 2004. In vivo substrate specificity of periplasmic disulfide oxidoreductases. *J Biol Chem* **279:** 12967-12973.
Hiramatsu, T., Guo, Y., and Hosoya, T. 2007. 3-Azidodifluoromethyl-3H-diazirin-3-yl group as an all-in-one functional group for radioisotope-free photoaffinity labeling. *Organic & biomolecular chemistry* **5:** 2916-2919.
Holzapfel, E., Eisner, G., Alami, M., Barrett, C.M., Buchanan, G., Luke, I., Betton, J.M., Robinson, C., Palmer, T., Moser, M., et al. 2007. The entire N-terminal half of TatC is involved in twin-arginine precursor binding. *Biochemistry* **46:** 2892-2898.
Hu, J.C. 2001. Model systems: Studying molecular recognition using bacterial n-hybrid systems. *Trends Microbiol* **9:** 219-222.
Huang, D., and Shusta, E.V. 2005. Secretion and surface display of green fluorescent protein using the yeast Saccharomyces cerevisiae. *Biotechnol Prog* **21:** 349-357.
Huber, D., Boyd, D., Xia, Y., Olma, M.H., Gerstein, M., and Beckwith, J. 2005a. Use of thioredoxin as a reporter to identify a subset of Escherichia coli signal sequences that promote signal recognition particle-dependent translocation. *J Bacteriol* **187:** 2983-2991.
Huber, D., Cha, M.I., Debarbieux, L., Planson, A.G., Cruz, N., Lopez, G., Tasayco, M.L., Chaffotte, A., and Beckwith, J. 2005b. A selection for mutants that interfere with folding of Escherichia coli thioredoxin-1 in vivo. *Proc Natl Acad Sci U S A* **102:** 18872-18877.
Ilbert, M., Mejean, V., Giudici-Orticoni, M.T., Samama, J.P., and Iobbi-Nivol, C. 2003. Involvement of a mate chaperone (TorD) in the maturation pathway of molybdoenzyme TorA. *J Biol Chem* **278:** 28787-28792.
Ize, B., Gerard, F., Zhang, M., Chanal, A., Voulhoux, R., Palmer, T., Filloux, A., and Wu, L.F. 2002. In vivo dissection of the Tat translocation pathway in Escherichia coli. *J Mol Biol* **317:** 327-335.
Jack, R.L., Sargent, F., Berks, B.C., Sawers, G., and Palmer, T. 2001. Constitutive expression of Escherichia coli tat genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth. *J Bacteriol* **183:** 1801-1804.
Jongbloed, J.D., Antelmann, H., Hecker, M., Nijland, R., Bron, S., Airaksinen, U., Pries, F., Quax, W.J., van Dijl, J.M., and Braun, P.G. 2002. Selective contribution of the twin-arginine translocation pathway to protein secretion in Bacillus subtilis. *J Biol Chem* **277:** 44068-44078.
Jormakka, M., Tornroth, S., Byrne, B., and Iwata, S. 2002. Molecular basis of proton motive force generation: structure of formate dehydrogenase-N. *Science (New York, N.Y* **295:** 1863-1868.
Kadokura, H., Katzen, F., and Beckwith, J. 2003. Protein disulfide bond formation in prokaryotes. *Annu Rev Biochem* **72:** 111-135.
Kaether, C., and Gerdes, H.H. 1995. Visualization of protein transport along the secretory pathway using green fluorescent protein. *FEBS Lett* **369:** 267-271.
Kapust, R.B., and Waugh, D.S. 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8:** 1668-1674.

Karimova, G., Dautin, N., and Ladant, D. 2005. Interaction network among Escherichia coli membrane proteins involved in cell division as revealed by bacterial two-hybrid analysis. *J Bacteriol* **187:** 2233-2243.
Karimova, G., Ladant, D., and Ullmann, A. 2002. Two-hybrid systems and their usage in infection biology. *Int J Med Microbiol* **292:** 17-25.
Karzai, A.W., Roche, E.D., and Sauer, R.T. 2000. The SsrA-SmpB system for protein tagging, directed degradation and ribosome rescue. *Nat Struct Biol* **7:** 449-455.
Ki, J.J., Kawarasaki, Y., Gam, J., Harvey, B.R., Iverson, B.L., and Georgiou, G. 2004. A periplasmic fluorescent reporter protein and its application in high-throughput membrane protein topology analysis. *J Mol Biol* **341:** 901-909.
Kim, J., Rusch, S., Luirink, J., and Kendall, D.A. 2001. Is Ffh required for export of secretory proteins? *FEBS Lett* **505:** 245-248.
Kipping, M., Lilie, H., Lindenstrauss, U., Andreesen, J.R., Griesinger, C., Carlomagno, T., and Bruser, T. 2003. Structural studies on a twin-arginine signal sequence. *FEBS Lett* **550:** 18-22.
Kouzarides, T., and Ziff, E. 1988. The role of the leucine zipper in the fos-jun interaction. *Nature* **336:** 646-651.
Krebber, A., Bornhauser, S., Burmester, J., Honegger, A., Willuda, J., Bosshard, H.R., and Pluckthun, A. 1997. Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *Journal of immunological methods* **201:** 35-55.
Kreutzenbeck, P., Kroger, C., Lausberg, F., Blaudeck, N., Sprenger, G.A., and Freudl, R. 2007a. Escherichia coli twin-arginine (Tat) mutant translocases possessing relaxed signal peptide recognition specificities. *J Biol Chem*.
Kreutzenbeck, P., Kroger, C., Lausberg, F., Blaudeck, N., Sprenger, G.A., and Freudl, R. 2007b. Escherichia coli twin arginine (Tat) mutant translocases possessing relaxed signal peptide recognition specificities. *J Biol Chem* **282:** 7903-7911.
Kumamoto, C.A. 1991. Molecular chaperones and protein translocation across the Escherichia coli inner membrane. *Mol Microbiol* **5:** 19-22.
Lang, B.F., Burger, G., O'Kelly, C.J., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., and Gray, M.W. 1997. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* **387:** 493-497.
Lavander, M., Ericsson, S.K., Broms, J.E., and Forsberg, A. 2007. Twin arginine translocation in Yersinia. *Advances in experimental medicine and biology* **603:** 258-267.
Lee, P.A., Buchanan, G., Stanley, N.R., Berks, B.C., and Palmer, T. 2002. Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation. *J Bacteriol* **184:** 5871-5879.
Lee, P.A., Orriss, G.L., Buchanan, G., Greene, N.P., Bond, P.J., Punginelli, C., Jack, R.L., Sansom, M.S., Berks, B.C., and Palmer, T. 2006. Cysteine-scanning mutagenesis and disulfide mapping studies of the conserved domain of the twin-arginine translocase TatB component. *J Biol Chem* **281:** 34072-34085.
Lequette, Y., Odberg-Ferragut, C., Bohin, J.P., and Lacroix, J.M. 2004. Identification of mdoD, an mdoG paralog which encodes a twin-arginine-dependent periplasmic

protein that controls osmoregulated periplasmic glucan backbone structures. *J Bacteriol* **186:** 3695-3702.
Levchenko, I., Seidel, M., Sauer, R.T., and Baker, T.A. 2000. A specificity-enhancing factor for the ClpXP degradation machine. *Science* **289:** 2354-2356.
Levy, R., Weiss, R., Chen, G., Iverson, B.L., and Georgiou, G. 2001. Production of correctly folded Fab antibody fragment in the cytoplasm of Escherichia coli trxB gor mutants via the coexpression of molecular chaperones. *Protein Expr Purif* **23:** 338-347.
Link, A.J., Mock, M.L., and Tirrell, D.A. 2003. Non-canonical amino acids in protein engineering. *Curr Opin Biotechnol* **14:** 603-609.
Link, A.J., Vink, M.K., and Tirrell, D.A. 2007a. Preparation of the functionalizable methionine surrogate azidohomoalanine via copper-catalyzed diazo transfer. *Nat Protoc* **2:** 1879-1883.
Link, A.J., Vink, M.K., and Tirrell, D.A. 2007b. Synthesis of the functionalizable methionine surrogate azidohomoalanine using Boc-homoserine as precursor. *Nat Protoc* **2:** 1884-1887.
Luirink, J., High, S., Wood, H., Giner, A., Tollervey, D., and Dobberstein, B. 1992. Signal-sequence recognition by an Escherichia coli ribonucleoprotein complex. *Nature* **359:** 741-743.
Luirink, J., von Heijne, G., Houben, E., and de Gier, J.W. 2005. Biogenesis of inner membrane proteins in Escherichia coli. *Annual review of microbiology* **59:** 329-355.
Macinga, D.R., and Rather, P.N. 1996. aarD, a Providencia stuartii homologue of cydD: role in 2'-N-acetyltransferase expression, cell morphology and growth in the presence of an extracellular factor. *Mol Microbiol* **19:** 511-520.
Maillard, J., Spronk, C.A., Buchanan, G., Lyall, V., Richardson, D.J., Palmer, T., Vuister, G.W., and Sargent, F. 2007. Structural diversity in twin-arginine signal peptide-binding proteins. *Proc Natl Acad Sci U S A* **104:** 15641-15646.
Malik, P., Terry, T.D., Bellintani, F., and Perham, R.N. 1998. Factors limiting display of foreign peptides on the major coat protein of filamentous bacteriophage capsids and a potential role for leader peptidase. *FEBS Lett* **436:** 263-266.
Mangels, D., Mathers, J., Bolhuis, A., and Robinson, C. 2005. The core TatABC complex of the twin-arginine translocase in Escherichia coli: TatC drives assembly whereas TatA is essential for stability. *J Mol Biol* **345:** 415-423.
Manoil, C., Mekalanos, J.J., and Beckwith, J. 1990. Alkaline phosphatase fusions: sensors of subcellular location. *J Bacteriol* **172:** 515-518.
Martin, J.F. 2004. Phosphate control of the biosynthesis of antibiotics and other secondary metabolites is mediated by the PhoR-PhoP system: an unfinished story. *J Bacteriol* **186:** 5197-5201.
McDevitt, C.A., Buchanan, G., Sargent, F., Palmer, T., and Berks, B.C. 2006. Subunit composition and in vivo substrate-binding characteristics of Escherichia coli Tat protein complexes expressed at native levels. *Febs J* **273:** 5656-5668.
Mejean, V., Iobbi-Nivol, C., Lepelletier, M., Giordano, G., Chippaux, M., and Pascal, M.C. 1994. TMAO anaerobic respiration in Escherichia coli: involvement of the tor operon. *Mol Microbiol* **11:** 1169-1179.

Muller, S.D., De Angelis, A.A., Walther, T.H., Grage, S.L., Lange, C., Opella, S.J., and Ulrich, A.S. 2007. Structural characterization of the pore forming protein TatA(d) of the twin-arginine translocase in membranes by solid-state (15)N-NMR. *Biochim Biophys Acta*.

Nakamoto, H., and Bardwell, J.C. 2004. Catalysis of disulfide bond formation and isomerization in the Escherichia coli periplasm. *Biochim Biophys Acta* **1694:** 111-119.

Ochsner, U.A., Snyder, A., Vasil, A.I., and Vasil, M.L. 2002. Effects of the twin-arginine translocase on secretion of virulence factors, stress response, and pathogenesis. *Proc Natl Acad Sci U S A* **99:** 8312-8317.

Oresnik, I.J., Ladner, C.L., and Turner, R.J. 2001. Identification of a twin-arginine leader-binding protein. *Mol Microbiol* **40:** 323-331.

Orriss, G.L., Tarry, M.J., Ize, B., Sargent, F., Lea, S.M., Palmer, T., and Berks, B.C. 2007. TatBC, TatB, and TatC form structurally autonomous units within the twin arginine protein transport system of Escherichia coli. *FEBS Lett* **581:** 4091-4097.

Ostermeier, M. 2003. Synthetic gene libraries: in search of the optimal diversity. *Trends in biotechnology* **21:** 244-247.

Oyama, R., Takashima, H., Yonezawa, M., Doi, N., Miyamoto-Sato, E., Kinjo, M., and Yanagawa, H. 2006. Protein-protein interaction analysis by C-terminally specific fluorescence labeling and fluorescence cross-correlation spectroscopy. *Nucleic Acids Res* **34:** e102.

Paetzel, M., Karla, A., Strynadka, N.C., and Dalbey, R.E. 2002. Signal peptidases. *Chemical reviews* **102:** 4549-4580.

Paleos, C.M., Tsiourvas, D., and Sideratou, Z. 2004. Hydrogen bonding interactions of liposomes simulating cell-cell recognition. A review. *Orig Life Evol Biosph* **34:** 195-213.

Palmer, T., Sargent, F., and Berks, B.C. 2004. Light traffic: photo-crosslinking a novel transport system. *Trends Biochem Sci* **29:** 55-57.

Palmer, T., Sargent, F., and Berks, B.C. 2005. Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol* **13:** 175-180.

Paradkar, A.S., Petrich, A.K., Leskiw, B.K., Aidoo, K.A., and Jensen, S.E. 1994. Transcriptional analysis and heterologous expression of the gene encoding beta-lactamase inhibitor protein (BLIP) from Streptomyces clavuligerus. *Gene* **144:** 31-36.

Paramban, R.I., Bugos, R.C., and Su, W.W. 2004. Engineering green fluorescent protein as a dual functional tag. *Biotechnology and bioengineering* **86:** 687-697.

Paschke, M., and Hohne, W. 2005. A twin-arginine translocation (Tat)-mediated phage display system. *Gene* **350:** 79-88.

Paschke, M., Tiede, C., and Hohne, W. 2007. Engineering a circularly permuted GFP scaffold for peptide presentation. *J Mol Recognit*.

Patterson, G.H., Knobel, S.M., Sharif, W.D., Kain, S.R., and Piston, D.W. 1997. Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy. *Biophysical journal* **73:** 2782-2790.

Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. 2006. Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* **24:** 79-88.

Perez-Rodriguez, R., Fisher, A.C., Perlmutter, J.D., Hicks, M.G., Chanal, A., Santini, C.L., Wu, L.F., Palmer, T., and Delisa, M.P. 2007. An Essential Role for the DnaK Molecular Chaperone in Stabilizing Over-expressed Substrate Proteins of the Bacterial Twin-arginine Translocation Pathway. *J Mol Biol* **367:** 715-730.

Pommier, J., Mejean, V., Giordano, G., and Iobbi-Nivol, C. 1998. TorD, a cytoplasmic chaperone that interacts with the unfolded trimethylamine N-oxide reductase enzyme (TorA) in Escherichia coli. *J Biol Chem* **273:** 16615-16620.

Pop, O., Martin, U., Abel, C., and Muller, J.P. 2002. The twin-arginine signal peptide of PhoD and the TatAd/Cd proteins of Bacillus subtilis form an autonomous Tat translocation system. *J Biol Chem* **277:** 3268-3273.

Porcelli, I., de Leeuw, E., Wallis, R., van den Brink-van der Laan, E., de Kruijff, B., Wallace, B.A., Palmer, T., and Berks, B.C. 2002. Characterization and membrane assembly of the TatA component of the Escherichia coli twin-arginine protein transport system. *Biochemistry* **41:** 13690-13697.

Pradel, N., Ye, C., Livrelli, V., Xu, J., Joly, B., and Wu, L.F. 2003. Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic Escherichia coli O157:H7. *Infection and immunity* **71:** 4908-4916.

Priyadarshini, R., Popham, D.L., and Young, K.D. 2006. Daughter cell separation by penicillin-binding proteins and peptidoglycan amidases in Escherichia coli. *J Bacteriol* **188:** 5345-5355.

Punginelli, C., Maldonado, B., Grahl, S., Jack, R., Alami, M., Schroder, J., Berks, B.C., and Palmer, T. 2007. Cysteine scanning mutagenesis and topological mapping of the Escherichia coli twin-arginine translocase TatC Component. *J Bacteriol* **189:** 5482-5494.

Qiu, J., Swartz, J.R., and Georgiou, G. 1998. Expression of active human tissue-type plasminogen activator in Escherichia coli. *Applied and environmental microbiology* **64:** 4891-4896.

Randall, L.L., Crane, J.M., Lilly, A.A., Liu, G., Mao, C., Patel, C.N., and Hardy, S.J. 2005. Asymmetric binding between SecA and SecB two symmetric proteins: implications for function in export. *J Mol Biol* **348:** 479-489.

Rapoza, M.P., and Webster, R.E. 1993. The filamentous bacteriophage assembly proteins require the bacterial SecA protein for correct localization to the membrane. *J Bacteriol* **175:** 1856-1859.

Rapp, M., Seppala, S., Granseth, E., and von Heijne, G. 2007. Emulating membrane protein evolution by rational design. *Science* **315:** 1282-1284.

Rather, P.N., and Orosz, E. 1994. Characterization of aarA, a pleiotrophic negative regulator of the 2'-N-acetyltransferase in Providencia stuartii. *J Bacteriol* **176:** 5140-5144.

Rather, P.N., Parojcic, M.M., and Paradise, M.R. 1997. An extracellular factor regulating expression of the chromosomal aminoglycoside 2'-N-acetyltransferase of Providencia stuartii. *Antimicrobial agents and chemotherapy* **41:** 1749-1754.

Ribnicky, B., Van Blarcom, T., and Georgiou, G. 2007. A scFv antibody mutant isolated in a genetic screen for improved export via the twin arginine transporter pathway exhibits faster folding. *J Mol Biol* **369:** 631-639.

Richter, S., Lindenstrauss, U., Lucke, C., Bayliss, R., and Bruser, T. 2007. Functional Tat transport of unstructured, small, hydrophilic proteins. *J Biol Chem*.

Rietsch, A., Bessette, P., Georgiou, G., and Beckwith, J. 1997. Reduction of the periplasmic disulfide bond isomerase, DsbC, occurs by passage of electrons from cytoplasmic thioredoxin. *J Bacteriol* **179:** 6602-6608.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. 1999. A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol* **17:** 1030-1032.

Rizzuto, R., Brini, M., Pizzo, P., Murgia, M., and Pozzan, T. 1995. Chimeric green fluorescent protein as a tool for visualizing subcellular organelles in living cells. *Curr Biol* **5:** 635-642.

Robinson, C., and Bolhuis, A. 2001. Protein targeting by the twin-arginine translocation pathway. *Nat Rev Mol Cell Biol* **2:** 350-356.

Rodrigue, A., Chanal, A., Beck, K., Muller, M., and Wu, L.F. 1999. Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. *J Biol Chem* **274:** 13223-13228.

Rossier, O., and Cianciotto, N.P. 2005. The Legionella pneumophila tatB gene facilitates secretion of phospholipase C, growth under iron-limiting conditions, and intracellular infection. *Infection and immunity* **73:** 2020-2032.

Samaluru, H., Saisree, L., and Reddy, M. 2007. Role of SufI (FtsP) in cell division of Escherichia coli: evidence for its involvement in stabilization of the division assembly. *J Bacteriol*.

Samuelson, J.C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G.J., and Dalbey, R.E. 2000. YidC mediates membrane protein insertion in bacteria. *Nature* **406:** 637-641.

Santini, C.L., Bernadac, A., Zhang, M., Chanal, A., Ize, B., Blanco, C., and Wu, L.F. 2001. Translocation of jellyfish green fluorescent protein via the Tat system of Escherichia coli and change of its periplasmic localization in response to osmotic up-shock. *J Biol Chem* **276:** 8159-8164.

Santini, C.L., Ize, B., Chanal, A., Muller, M., Giordano, G., and Wu, L.F. 1998. A novel sec-independent periplasmic protein translocation pathway in Escherichia coli. *Embo J* **17:** 101-112.

Sargent, F., Berks, B.C., and Palmer, T. 2006. Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins. *FEMS Microbiol Lett* **254:** 198-207.

Sargent, F., Bogsch, E.G., Stanley, N.R., Wexler, M., Robinson, C., Berks, B.C., and Palmer, T. 1998. Overlapping functions of components of a bacterial Sec-independent protein export pathway. *Embo J* **17:** 3640-3650.

Sawers, R.G., Ballantine, S.P., and Boxer, D.H. 1985. Differential expression of hydrogenase isoenzymes in Escherichia coli K-12: evidence for a third isoenzyme. *J Bacteriol* **164:** 1324-1331.

Schatz, G., and Dobberstein, B. 1996. Common principles of protein translocation across membranes. *Science* **271:** 1519-1526.
Schiebel, E., Driessen, A.J., Hartl, F.U., and Wickner, W. 1991. Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell* **64:** 927-939.
Schierle, C.F., Berkmen, M., Huber, D., Kumamoto, C., Boyd, D., and Beckwith, J. 2003. The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway. *J Bacteriol* **185:** 5706-5713.
Scott, J.K., and Smith, G.P. 1990. Searching for peptide ligands with an epitope library. *Science* **249:** 386-390.
Shanmugham, A., Wong Fong Sang, H.W., Bollen, Y.J., and Lill, H. 2006. Membrane binding of twin arginine preproteins as an early step in translocation. *Biochemistry* **45:** 2243-2249.
Siemering, K.R., Golbik, R., Sever, R., and Haseloff, J. 1996. Mutations that suppress the thermosensitivity of green fluorescent protein. *Curr Biol* **6:** 1653-1663.
Silvestro, A., Pommier, J., and Giordano, G. 1988. The inducible trimethylamine-N-oxide reductase of Escherichia coli K12: biochemical and immunological studies. *Biochim Biophys Acta* **954:** 1-13.
Smith, G.P. 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228:** 1315-1317.
Stanley, N.R., Palmer, T., and Berks, B.C. 2000. The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in Escherichia coli. *J Biol Chem* **275:** 11591-11596.
Steiner, D., Forrer, P., Stumpp, M.T., and Pluckthun, A. 2006. Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nat Biotechnol* **24:** 823-831.
Stevenson, L.G., Strisovsky, K., Clemmer, K.M., Bhatt, S., Freeman, M., and Rather, P.N. 2007. Rhomboid protease AarA mediates quorum-sensing in Providencia stuartii by activating TatA of the twin-arginine translocase. *Proc Natl Acad Sci U S A* **104:** 1003-1008.
Strauch, E.M., and Georgiou, G. 2007a. A bacterial two-hybrid system based on the twin-arginine transporter pathway of E. coli. *Protein Sci* **16:** 1001-1008.
Strauch, E.M., and Georgiou, G. 2007b. Escherichia coli tatC Mutations that Suppress Defective Twin-Arginine Transporter Signal Peptides. *J Mol Biol*.
Subramanian, K., and Meyer, T. 1997. Calcium-induced restructuring of nuclear envelope and endoplasmic reticulum calcium stores. *Cell* **89:** 963-971.
Tam, R., and Saier, M.H., Jr. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiological reviews* **57:** 320-346.
Tokumoto, U., Nomura, S., Minami, Y., Mihara, H., Kato, S., Kurihara, T., Esaki, N., Kanazawa, H., Matsubara, H., and Takahashi, Y. 2002. Network of protein-protein interactions among iron-sulfur cluster assembly proteins in Escherichia coli. *Journal of biochemistry* **131:** 713-719.

Tree, J.J., Kidd, S.P., Jennings, M.P., and McEwan, A.G. 2005. Copper sensitivity of cueO mutants of Escherichia coli K-12 and the biochemical suppression of this phenotype. *Biochem Biophys Res Commun* **328:** 1205-1210.

Tu, G.F., Reid, G.E., Zhang, J.G., Moritz, R.L., and Simpson, R.J. 1995. C-terminal extension of truncated recombinant proteins in Escherichia coli with a 10Sa RNA decapeptide. *J Biol Chem* **270:** 9322-9326.

Tullman-Ercek, D., DeLisa, M.P., Kawarasaki, Y., Iranpour, P., Ribnicky, B., Palmer, T., and Georgiou, G. 2007. Export pathway selectivity of Escherichia coli twin arginine translocation signal peptides. *J Biol Chem* **282:** 8309-8316.

Turner, R.J., Papish, A.L., and Sargent, F. 2004. Sequence analysis of bacterial redox enzyme maturation proteins (REMPs). *Can J Microbiol* **50:** 225-238.

van der Wolk, J.P., de Wit, J.G., and Driessen, A.J. 1997. The catalytic cycle of the escherichia coli SecA ATPase comprises two distinct preprotein translocation events. *Embo J* **16:** 7297-7304.

Walsh, G. 2005. Biopharmaceuticals: recent approvals and likely directions. *Trends in biotechnology* **23:** 553-558.

Wexler, M., Sargent, F., Jack, R.L., Stanley, N.R., Bogsch, E.G., Robinson, C., Berks, B.C., and Palmer, T. 2000. TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export. *J Biol Chem* **275:** 16717-16722.

Widdick, D.A., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschroder, M., and Palmer, T. 2006. The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. *Proc Natl Acad Sci U S A* **103:** 17927-17932.

Yamada, K., Sanzen, I., Ohkura, T., Okamoto, A., Torii, K., Hasegawa, T., and Ohta, M. 2007. Analysis of twin-arginine translocation pathway homologue in Staphylococcus aureus. *Current microbiology* **55:** 14-19.

Yen, M.R., Tseng, Y.H., Nguyen, E.H., Wu, L.F., and Saier, M.H., Jr. 2002. Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system. *Arch Microbiol* **177:** 441-450.

Zeytun, A., Jeromin, A., Scalettar, B.A., Waldo, G.S., and Bradbury, A.R. 2003. Fluorobodies combine GFP fluorescence with the binding characteristics of antibodies. *Nat Biotechnol* **21:** 1473-1479.

Zeytun, A., Jeromin, A., Scalettar, B.A., Waldo, G.S., and Bradbury, A.R. 2004. Retraction. Fluorobodies combine GFP fluorescence with the binding characteristics of antibodies. *Nat Biotechnol* **22:** 601.

Vita

Eva-Maria Strauch, the daughter of Waltraud Strauch and Rainer Strauch, was born in Kassel (Germany) on March 27, 1979. After completing her work at the German high school Friedrichsgymnasium in Kassel in 1998, she entered the Julius-Maximilians Universität in Würzburg to study Biology. She majored in Biochemistry with a minor in Microbiology, Genetics and Biotechnology.

In 2002, she transferred to the University of Texas at Austin, entering the Graduate School of Chemistry and Biochemistry where she earned her Master's degree in Biochemistry 2005 and is now working towards her Doctorate in Biochemistry.

Permanent address: Am Kammerberg 99, 34292 Ahnatal, Germany

This dissertation was typed by the author